Spring 3-10-2014

Hedonics, Reward Prediction, and Reinforcement Learning in Schizophrenia: Relationships to Anhedonia and Avolition

Erin Dowd
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences
Neurosciences

Dissertation Examination Committee:
Deanna Barch, Chair
Kevin Black
Todd Braver
Tamara Hershey
Jose Mathews
Steven Petersen

Hedonics, Reward Prediction, and Reinforcement Learning in Schizophrenia: Relationships to Anhedonia and Avolition

by

Erin Connor Dowd

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2014

St. Louis, Missouri

Table of Contents

Acknowledgements

First and foremost, I would like to thank my advisor, Dr. Deanna Barch. Deanna's commitment to her students involves not only teaching them to conduct sound, clinically meaningful science, but training them to be successful well-rounded academics as well. I am indebted to her for challenging me, for giving me freedom to develop my own questions and thought processes, for providing external motivation (read: deadlines) when needed, for ensuring that I make contacts in the field, and for guidance and advice about data, careers, and life. It has been an honor and a pleasure to train in her lab.

I would also like to thank the members of my thesis advisory committee for taking time out of their busy schedules to provide advice and feedback on my work. I am grateful to Todd Braver for his encouragement and guidance on the computational aspects of my project, and to Joel Perlmutter for sharing his expertise to help me refine my experimental question and hypotheses. Kevin Black's advice was invaluable in keeping sight of the clinical implications of my work. Steve Petersen's inspiring example of disciplined, open-minded consideration of the data is a model that I will carry throughout my career. I am also grateful to Tammy Hershey and Jose Mathews for joining the committee for my thesis exam, and look forward to hearing their feedback and advice for future work.

There are also several researchers outside of Washington University who have offered me valuable advice on the work here. I would like to thank Michael Frank for taking the time to consult with me at several stages throughout my project, Anne Collins for her help in the practical aspects of implementing the computational model, and Jim Waltz for giving thoughtful, in-depth feedback on several iterations of my work.

I also wish to thank the other members of the Barch and Braver labs for their insight, enthusiasm, and camaraderie. In addition, I am indebted to Melissa Cornejo and Emily Thomason, without whom I would still be collecting data, and Carol Cox for sharing her hard-earned knowledge to help things go smoothly.

I would also like to express my deepest appreciation for my parents, John and Janee Connor, who taught me to pursue my fullest potential unapologetically, but with honesty, kindness, and joy. No words can express the extent to which I am indebted to them for these lessons.

Finally, I want to thank my husband, Michael Dowd, whose unconditional support throughout this process has been nothing short of astonishing. He has played every role from consulting on my research to shouldering more than his fair share of domestic duties. He is a remarkable thinker who challenges me daily, and his influence on my work, both directly and indirectly, is immeasurable.

This thesis is dedicated to the memory of my grandfather, Bill Keary, my role model of intellectual curiosity and disciplined thought. His infectious love of learning was an inspiration to all who knew him.

This work was supported financially by an NSF IGERT program fellowship, an NRSA F30 fellowship from NIMH, and by R01 MH071616.

ABSTRACT OF THE DISSERTATION

Hedonics, Reward Prediction, and Reinforcement Learning in Schizophrenia: Relationships to Anhedonia and Avolition

by

Erin Connor Dowd

Doctor of Philosophy in Biology & Biomedical Sciences

Neurosciences

Washington University in St. Louis, 2014

Professor Deanna M. Barch, Chairperson

Anhedonia (a reduced experience of pleasure) and avolition (a reduction in goal-directed activity) are common features of schizophrenia that have substantial effects on functional outcome, but are poorly understood and treated. Here, we examine the potential contributions of three processes involved in the translation of reward information in the environment into goal-directed action among medicated individuals with schizophrenia. The first process, hedonics, or the enjoyment of pleasant experiences, was examined using an fMRI study of emotional experience in response to affective stimuli. Results were similar between patients and controls at a group level, but patients with greater anhedonia/avolition demonstrated reduced self-reports and brain activity in response to positive stimuli in ventral striatum and amygdala, regions associated with salience and reward. The second process we examined was reward prediction, which describes anticipatory responses to cues that predict reward. This process is thought to be mediated by the mesolimbic dopamine system, which shows evidence of dysregulation in schizophrenia. We examined reward prediction using a Pavlovian paradigm with monetary reward, using fMRI to examine brain responses during reward anticipation and receipt. Responses to reward receipt were largely intact in schizophrenia, while anticipation responses in ventral striatum, a major target of dopaminergic afferents, were reduced in those patients who were higher in anhedonia/avolition. The final process we examined was reinforcement learning, or the process by which positive and negative

feedback influences trial-and-error choice behavior. Participants underwent fMRI during a reinforcement learning task with probabilistic feedback. Evidence from behavior and computational modeling suggested impairment in learning from positive feedback. However, neuroimaging data revealed largely intact striatal activation during both choice and feedback, a surprising result given the role of dopaminergic influence on corticostriatal circuits in mediating reinforcement learning. Instead, there was some evidence for reduced cortical responses to choice execution among patients, and to positive feedback among those patients higher in anhedonia/avolition. Together, these studies suggest that impairments in hedonics, reward prediction, and reinforcement learning may play a role in anhedonia/avolition in schizophrenia, but that these impairments are not sufficient causes of these symptoms; impairments in other higher-level cognitive processes are also likely to contribute.

Chapter 1.

Reward Processes and Negative Symptoms in Schizophrenia

Portions of this chapter were published in the September, 2010 issue of Schizophrenia Bulletin, to which I contributed conceptual input, literature review, and writing of some sections.

Reference: Barch DM & Dowd EC. Goal representations and motivational drive in schizophrenia: the role of prefrontal-striatal interactions. Schizophrenia Bulletin (2010) vol. 36 (5) pp. 919-34.

Schizophrenia and Negative Symptoms

Schizophrenia is a severe psychiatric illness with a complex and heterogeneous presentation. Among neuropsychiatric illnesses, schizophrenia is one of the most disabling, and renders heavy personal, economic, and social burdens (Knapp et al 2004) that place it among the top ten causes of long-term disability worldwide (Mueser and McGurk 2004). While schizophrenia is most commonly associated with positive symptoms such as hallucinations and delusions, clinicians have long recognized that cognitive deficits and negative symptoms – such as affective flattening and reduced motivation – are important symptoms that contribute substantially to disability and poor quality of life in this illness (Kirkpatrick et al 2006; Leifker et al 2009; Milev et al 2005; Nuechterlein et al 2012). In fact, negative symptoms may have a greater impact on functional outcome than positive symptoms (Rabinowitz et al 2012), but currently available treatments are often ineffective in reducing them (Kirkpatrick et al 2006). As revealed by factor-analysis studies of negative symptom scales such as the Scale for the Assessment of Negative Symptoms (SANS), negative symptoms consist of two major factors: expressive deficits, and anhedonia/avolition (Messinger et al 2011). The expressive deficits factor includes impairments in emotional expression such as flat affect and alogia (poverty of speech), and the anhedonia/avolition factor, on which I focus here, includes anhedonia (a reduction in the ability to experience pleasure) and avolition (a reduction in the motivation to initiate or persist in goal-directed behavior). The work described here strives to examine the anhedonia/avolition symptom domain by using a perspective grounded in affective and computational neuroscience to identify and systematically examine several potential contributors to these symptoms among individuals with schizophrenia.

In order to study the etiology of anhedonia and avolition in schizophrenia, it is helpful to outline several component processes that are required for the translation of information about rewarding experiences into behavioral responses designed to obtain these rewards again in the future (Wallis 2007). These processes represent major elements of reward processing and decision making as identified in basic animal and human studies (Berridge 2007; Wallis 2007), and while they are certainly an oversimplification, these processes provide a useful set of organizing principles for studying reductions in goal-directed behavior among individuals with schizophrenia. The first of these components is ***hedonics***

or "liking", which reflects the enjoyment that results from experiencing a pleasurable stimulus or event. Impairments in hedonics/"liking" would yield anhedonia by definition, but may also yield avolition if the reduction in pleasure causes a lack of motivation to seek out similar events in the future. The possibility that hedonic processes are impaired in schizophrenia is examined in Chapter 2: Anhedonia and Emotional Experience in Schizophrenia. The second component process is ***reward prediction and "wanting"***, which describes the ability of cues that predict rewards to trigger anticipatory responses in expectation of reward. Impairment in this system would be expected to yield avolition by reducing the salience of reward-predictive cues in the environment, thus reducing the motivational drive to seek out the associated rewards. I ask whether reward prediction is impaired in schizophrenia in Chapter 3: Pavlovian Reward Prediction in Schizophrenia. The third component, which is closely related to reward prediction and wanting, is ***reinforcement learning***, or the conditioning process by which choices are shaped by feedback over time to maximize reward. If this process is impaired, subjects would have difficulty making choices that are likely to yield rewards even if they are motivated to obtain the rewards and enjoy them when they occur. This possibility is examined in Chapter 4: Probabilistic Reinforcement Learning in Schizophrenia. In addition to these three components, which address the contribution of basic reward processing to goal-directed behavior, several additional computations must be performed on the reward information before a behavioral response is carried out. These include the integration and updating of value information, effort computation, cost-benefit analysis, and finally generation of an action plan to obtain a valued outcome. While these processes were not examined in the studies conducted here, impairment in any of these domains would also be expected to yield reductions in goal-directed behavior, a possibility I return to in Chapter 5: Conclusions and Future Directions.

Hedonics or "liking" in schizophrenia

Hedonic or "liking" responses to pleasurable stimuli were long thought to be mediated by the mesolimbic dopamine (DA) system, primarily due to evidence that lesions to this system decreased reward-seeking behaviors in rats (Wise 1982). Since this time, however, work by Berridge and colleagues has shown that dopamine depletion does not reduce pleasure responses that do not depend upon motivational drive, such as orofacial reactions to sweet tastes (Berridge 2004). Accordingly, more recent

work suggests that hedonic responses are mediated not by DA, but by opioid- and GABA-dependent mechanisms in the ventral striatum, amygdala, and orbitofrontal cortex (Burgdorf and Panksepp 2006), while DA is associated with the related, but dissociable, functions of "wanting" (Berridge and Kringelbach 2008) and learning (Schultz 2007).

In the schizophrenia literature, current laboratory data on hedonics is somewhat at odds with the historical clinical picture. Anhedonia has long been considered a core clinical feature of schizophrenia (Bleuler 1911; Kraepelin 1971), but a large body of literature has shown that individuals with schizophrenia report experiencing similar levels of pleasure as controls when describing their feelings "in-the-moment" (see (Cohen and Minor 2008; Kring and Moran 2008) for reviews). In numerous laboratory-based experiments of emotional experience, self-reported valence and arousal responses to emotion-eliciting stimuli such as pictures, words, faces, and sounds have been found to be similar between patients and controls. Similarly, naturalistic experience-sampling studies examining emotional responses to everyday events have shown that individuals with schizophrenia and controls experience comparable increases in positive emotion when engaging in pleasurable activities (Gard et al 2007; Oorschot et al 2011), although patients do experience less positive affect overall, perhaps due to having fewer pleasurable events in their every day lives (Myin-Germeys et al 2000). In addition to these self-report measures of emotion, intact hedonic responses in schizophrenia may also be inferred from the effects of positive emotion on other psychological processes. For example, startle responses are similarly reduced in both individuals with schizophrenia and controls following presentation of pleasant stimuli (Curtis et al 1999; Schlenker et al 1995), and memory is enhanced for positively valenced emotional stimuli among both patients and controls (Hall et al 2007; Mathews and Barch 2004); although one study did find a failure of memory enhancement for positive, but not negative, stimuli (Herbener et al 2007).

A similar pattern of intact responses to rewarding stimuli among patients with schizophrenia is also seen in neuroimaging studies that examine striatal responses to receipt of monetary reward in monetary incentive delay (MID) tasks. A number of such studies have reported intact ventral striatal responses in patients treated with either typical or atypical antipsychotics (Kirsch et al 2007; Schlagenhauf et al 2009; Simon et al 2010; Walter et al 2009). One of these studies did show reduced striatal responses to loss-avoidance among the individuals with schizophrenia, although group differences

in ventral striatal responses to reward receipt were not significant (Schlagenhauf et al 2009). However, while striatal responses to receipt of money seem to be largely intact in schizophrenia, some of these studies did report abnormal cortical responses. Specifically, reduced reward-related responses in patients as compared to controls have been reported in ventromedial prefrontal cortex (VMPFC) (Schlagenhauf et al 2009), rostral anterior cingulate cortex (rACC), inferior frontal gyrus, and amygdala (Waltz et al 2010), and reduced salience coding in has been reported in right ventrolateral prefrontal cortex (Walter et al 2010; Walter et al 2009).

Of note, a somewhat more mixed picture has arisen from functional neuroimaging studies that have examined brain responses to other types of pleasurable or rewarding stimuli in schizophrenia. A study of juice rewards revealed reduced activation upon juice delivery in a number of regions including bilateral putamen, insula, precentral gyrus, and postcentral gyrus (Waltz et al 2009). Studies using positive emotional pictures have revealed evidence for reduced ventral striatal activation in a mixed sample of medicated and unmedicated patients (Taylor et al 2005). In unmedicated patients, reduced responses to pleasant pictures have been identified in medial frontal cortex, dorsolateral prefrontal cortex (DLPFC), orbitofrontal cortex (OFC), and insula (Paradiso et al 2003). Studies of pleasant olfactory stimuli have revealed reduced activation in thalamus (Schneider et al 2007), insula, and OFC (Plailly et al 2006) among patients as compared to controls. However, a recent meta-analyses of studies of emotion perception and experience in schizophrenia failed to reveal evidence for systematic differences in activation patterns between patients and controls for positive stimuli (Taylor et al 2012), suggesting that the abnormal patterns of activity in these studies of emotional pictures and odors are not particularly robust. The studies of odors may also have been influenced by factors such as smoking status, which can alter olfactory and gustatory processing.

The abundance of evidence suggesting that hedonic experience may be intact in schizophrenia is puzzling given the long history of clinically documented anhedonia among these patients. Studies of anhedonia using self-report questionnaires nearly unequivocally report elevated anhedonia in patients as compared to controls (Trémeau 2006). The source of this discrepancy, referred to as the “emotion paradox” in schizophrenia, is as yet unresolved in the literature, but the most prominent hypotheses emphasize interactions between cognition and emotion. Some authors have suggested that the emotion

paradox may result from impairments in the ability to remember or generalize emotion experienced in the past, yielding the appearance of emotion deficits on self-report measures with retrospective, abstracted, or hypothetical foci (Burbridge and Barch 2007; Strauss and Gold 2012). Others (Gard et al 2007) have emphasized the differences between consummatory pleasure (hedonics) and anticipatory pleasure (pleasure resulting from reward prediction/wanting), pointing out that while these constructs have been shown to rely on physiologically dissociable systems, traditional self-report measures of anhedonia do not distinguish between them (Horan 2005). In an experience-sampling study designed to address this question, it was found that patients experienced significantly less anticipatory, but not consummatory, pleasure than controls, particularly for goal-directed activities (Gard et al 2007).

It is also important to note that given the heterogeneity of symptoms among individuals with a schizophrenia diagnosis, there may be important individual differences in anhedonia severity that relate to individual experiences of pleasure. In several studies of emotional stimuli including pictures, words, sounds, faces, and films, higher anhedonia scores were associated with reduced positive affect in response to positive stimuli among patients (Blanchard et al 1994), or with blunted emotional responses to both positive and negative stimuli in both patients and controls (Burbridge and Barch 2007; Herbener et al 2007). In the neuroimaging literature, there is evidence that the magnitude of the reduction in striatal responses to the receipt of juice among patients is associated with the severity of anhedonia scores (Waltz et al 2009). In addition, Walter and colleagues report that salience coding in ventrolateral prefrontal cortex, which is reduced in patients, is correlated with negative symptom severity (Walter et al 2009) or anhedonia (Walter et al 2010). Interestingly, one study found that the magnitude of the reward receipt response in the ventral striatum was inversely associated with severity of depression, but not with anhedonia (Simon et al 2010), raising questions about the specificity of these effects. This evidence suggests that hedonic experience and its associated BOLD activation may vary with individual differences in anhedonia; however, interpretation of these results is complicated by the fact that the “anhedonia” measures here may have been influenced by anticipatory and motivational processes as well as by hedonics per se.

In sum, the self-report literature provides fairly consistent evidence for intact in-the-moment hedonic experience in schizophrenia, though there is evidence that greater self-reports of anhedonia are

associated with reduced hedonic experience. In addition, the neuroimaging literature on the receipt of monetary rewards shows relatively consistent evidence for intact striatal responses to the receipt of such rewards. Cortical responses to reward receipt are more variable, particularly for rewarding or pleasant stimuli other than money. These studies have identified abnormal responses among patients in striatal regions as well as ventrolateral prefrontal cortex, insula, OFC, and VMPFC; however, the patterns identified have been inconsistent across studies. Further, studies that have examined individual differences in negative symptoms suggest an important relationship between anhedonia (or motivational disturbances that may influence anhedonia questionnaires) and self-reports of hedonic experience. There are also hints of a relationship between anhedonia and reduced reward responses in striatum and VLPFC, but too few studies have been appropriately powered for individual difference analyses to draw strong conclusions about these relationships. Chapter 2 aims to address some of the open questions in this domain by examining functional activation during the experience of several types of non-monetary emotional stimuli in schizophrenia, and its relationship to self-reported anhedonia, in a large sample of patients and community controls.

Reward Prediction and "Wanting" in Schizophrenia

A second process required for translating reward information into goal-directed behavior is reward prediction and "wanting". An extensive body of literature has revealed that this process is mediated by the midbrain dopamine system and its projections to dorsal and ventral striatum (Berridge 2004; Schultz 2007). Converging data from animal, human, and computational studies suggests that phasic dopaminergic activity in the ventral tegmental area and substantia nigra mediates the acquisition of stimulus-reward associations over repeated pairings. A prominent theory holds that this phasic dopaminergic activity codes reward prediction errors that serve as teaching signals, guiding the formation of associations between rewards and the cues that predict them (Schultz et al 1997). A closely related, though slightly different, theory emphasizes the role of the dopamine-mediated learning process in transferring incentive salience from the reward itself to reward-predicting cues, thus imbuing these cues with motivational properties themselves (i.e., a **"**wanting**"** response (Berridge 2004)). We do not aim to

adjudicate between these theories here, but instead point out that both involve a role of dopamine in establishing anticipatory and approach motivational responses to reward-predictive cues. In support of both theories, during the process of learning, dopaminergic neurons in the ventral tegmental area initially fire in response to unexpected rewards, but after repeated cue-reward pairings, begin to fire upon cue presentation instead (Schultz et al 1993). When a cued reward is then omitted, dopaminergic firing momentarily drops below its basal level. This behavior is consistent with a role of phasic dopaminergic activity in coding prediction errors, or the difference between the magnitudes of rewards that were expected, and those that were actually received. This difference may be used to update reward predictions on each trial until prediction is optimized, thereby driving learning. This iterative process is formalized in computational algorithms such as temporal difference models (Sutton 1988), whose simulations show considerable agreement with experimental data from dopaminergic neurons in animal studies (Montague et al 1996; Schultz 2007). Studies of reward prediction in humans have also shown functional activity consistent with dopaminergic coding of prediction errors (O'Doherty 2004). Studies examining BOLD responses to reward-predictive cues have revealed anticipatory activation in regions including dorsal and ventral striatum, which are major targets of dopaminergic projections from the midbrain (Knutson et al 2001; Knutson et al 2000). Further, during Pavlovian conditioning, regions including striatum and orbitofrontal cortex have shown evidence of activation that correlates with prediction errors derived from temporal difference models fit to the behavioral data (McClure et al 2003; O'Doherty et al 2003b).

In schizophrenia, there is robust evidence for altered striatal dopaminergic function, which raises a particular interest in reward prediction in this disorder. The specifics of dopaminergic pathophysiology in schizophrenia are as yet unknown, but recent meta-analyses of PET studies in schizophrenia indicate that striatal dopaminergic disturbances in schizophrenia are primarily presynaptic (Fusar-Poli and Meyer-Lindenberg 2012; Howes et al 2012). These studies have revealed that striatal dopamine synthesis capacity, as assessed with [^{18}F]/[^{11}C]-DOPA, is elevated in schizophrenia (Fusar-Poli and Meyer-Lindenberg 2012), as are dopamine release as measured with [^{11}C]raclopride or [^{123}I]IBZM binding following acute amphetamine challenge (Howes et al 2009; Howes et al 2012) and synaptic dopamine levels as assessed by raclopride binding before and after pharmacologic dopamine depletion (Howes et

al 2012; Kegeles et al 2010). These presynaptic abnormalities may result in spurious DA release (Howes et al 2009) that could increase the noise in the system, reducing the effectiveness of environmentally relevant phasic dopamine signals and yielding impaired reward prediction mechanisms.

In the fMRI literature, the main approach to studying reward prediction has been to examine reward anticipation responses, or activation in response to stimuli that have been associated with reward during pre-scan conditioning or explicit instruction. The monetary incentive delay (MID) task is commonly used in these studies, and involves presentation of a cue that indicates potential monetary gain or loss for the trial, followed by a reaction time task, then feedback on whether the money was won. In healthy adults, such studies have shown reward anticipation responses upon presentation of the incentive cues in regions including dorsal and ventral striatum (Knutson et al 2001; Knutson et al 2000). In individuals with schizophrenia, a number of studies using the MID task have reported reduced ventral striatal activation to reward-predicting cues among patients as compared to controls. This result has been seen in unmedicated individuals with schizophrenia (Juckel et al 2006b), as well as in individuals taking typical antipsychotics (Juckel et al 2006a), but not in individuals treated with atypicals (Juckel et al 2006a; Walter et al 2009; Waltz et al 2010), nor in prodromal individuals (Juckel et al 2012). Studies focusing on medication have shown reduced ventral striatal responses to incentive cues in individuals with schizophrenia taking typical compared to atypical antipsychotics (Kirsch et al 2007), as well as reductions in anticipatory ventral striatal responses in antipsychotic-naive patients that improved following treatment with amisulpride (Nielsen et al 2012). Reduced striatal responses to appetitive food cues have also been reported among individuals with schizophrenia as compared to controls (Grimm et al 2012).

Importantly, a number of these studies also showed a relationship between deficits in anticipatory ventral striatal activity and individual differences in negative symptom severity. In unmedicated (Juckel et al 2006b), typically medicated (Juckel et al 2006a), and primarily atypically medicated (Waltz et al 2010) patients, the severity of negative symptoms has been shown to predict the reduction in ventral striatal responses to anticipated monetary gains. Similarly, Simon and colleagues showed that the magnitude of the ventral striatal anticipatory response was inversely correlated with apathy (avolition) ratings among atypically medicated individuals with schizophrenia.

To summarize, a growing number of studies in the imaging literature suggests reduced ventral striatal reward prediction/"wanting" responses in unmedicated and typically medicated individuals with schizophrenia, while responses in patients taking atypical antipsychotics are more frequently intact. However, there is evidence that the magnitude of these striatal impairments may be related to the severity of negative symptoms even among atypically treated patients. It is important to note, however, that the vast majority of these results were derived from a single task. The MID task is a good choice for studies of reward prediction because of its consistently demonstrated ability to elicit robust striatal anticipatory responses, but it also requires subjects to execute speeded responses in order to obtain rewards, which may confound the ability to predict rewards with the ability to execute the actions required to earn them. Chapter 3 addresses this problem by examining reward prediction during a Pavlovian task, where anticipatory responses to reward-predictive cues are examined in a context wherein outcomes depend only on the cues presented and are not contingent upon the execution of a behavioral response.

Reinforcement Learning in Schizophrenia

The term Reinforcement Learning (RL) is derived from the artificial intelligence literature, and describes algorithms that solve trial-and-error learning problems by using positive and negative feedback to optimize choice behavior with the goal of maximizing reward rate. In the past two decades, interaction between the fields of neurophysiology, psychology and computational neuroscience has allowed the development of physiologically and behaviorally informed RL models, alongside neurophysiology and neuroimaging studies designed from information-processing standpoints influenced by RL (Cohen and Frank 2009; Dayan and Balleine 2002). Current data suggests that the same dopaminergic prediction error mechanisms underlying reward prediction and "wanting" also contribute to reinforcement learning. Phasic dopaminergic activity has been shown to modulate synaptic plasticity in the corticostriatal circuits associated with action selection (Calabresi et al 1997), thereby facilitating the strengthening of stimulus-response associations for actions that yield positive outcomes, and the weakening of those that do not (Frank et al 2004). Recent work in computational modeling suggests that the basal ganglia may be responsible for implicit forms of learning involving the gradual integration of reward information over

several trials, while top-down information from frontal cortex, OFC and amygdala is required for more rapid and explicit forms of learning such as those required in conditions of changing response contingencies (Frank and Claus 2006).

Despite the evidence of alterations in striatal dopamine function in schizophrenia, numerous studies have suggested that reinforcement learning is intact in this population when learning is relatively easy or implicit (Ceaser et al 2008; Elliott et al 1995; Heerey et al 2008; Hutton et al 1998; Jazbec et al 2007; Joyce et al 2002; Somlai et al 2011; Turner et al 2004; Tyson et al 2004; Weiler et al 2009), though with some exceptions (Oades 1997; Pantelis et al 1999). However, when the learning paradigms are more difficult and/or require the explicit use of representations about stimulus-reward contingencies, there is more consistent evidence of impaired reinforcement learning in patients (Gold et al 2012; Koch et al 2010; Morris et al 2008; Waltz et al 2007; Yilmaz et al 2012). This pattern has given rise to the theory that the striatum-mediated gradual reinforcement learning system may be intact in schizophrenia, while the more rapid, on-line, cortically-mediated learning systems are impaired (Gold et al 2008; Weickert et al 2002). In support of this theory, two studies of probabilistic reversal learning in schizophrenia showed that while reversal learning (which is thought to rely on rapid cortical mechanisms) was impaired, patients were able to acquire the initial probabilistic reward contingencies that may in part reflect striatally-mediated implicit learning (Waltz and Gold 2007; Weiler et al 2009). Additional evidence for this theory comes from the Weather Prediction Task, a probabilistic category-learning task on which performance is influenced by both implicit and explicit learning mechanisms Gluck, 2002 #145}. Several studies using this task have shown a relatively intact learning rate, but impaired asymptotic performance, in individuals with schizophrenia (Beninger et al 2003; Keri et al 2000; Keri et al 2005; Weickert et al 2002). This finding is consistent with the possibility that the normal learning curve reflects relatively intact striatal learning mechanisms, while the impaired overall performance reflects relatively impaired cortically-supported explicit learning mechanisms. However, more recent studies, including one with a much larger sample size than others in the literature (108 patients), found lower learning rates in patients than controls, suggesting impairments in striatal learning as well (Weickert et al 2010). There is some evidence that performance on this task may be influenced by medication status, with greater impairments in patients taking typical than atypical antipsychotics (Beninger et al 2003), though performance does not

appear to be influenced by antipsychotic dose (Weickert et al 2010). Other tasks have emphasized the difference between positive and negative reinforcement learning, and have yielded evidence that individuals with schizophrenia may have more difficulty learning from positive than negative feedback (Cheng et al 2012; Gold et al 2012; Waltz et al 2007), particularly if they are higher in negative symptoms (Gold et al 2012).

One major approach to studying reinforcement learning in human neuroimaging studies is to ask whether the pattern of functional activation in regions receiving dopamingeric projections is consistent with a prediction error signal. Broadly, this framework would predict an increase in striatal (presumably dopaminergically mediated) responses to unexpected rewards, and a decrease in striatal responses when predicted rewards do not occur. In a more nuanced approach, BOLD responses during learning tasks are analyzed to determine whether they correlate with prediction errors derived from a computational model fit to the behavioral data. Studies of this kind among healthy adults have revealed evidence of activation consistent with a prediction error function in dorsal and ventral striatum, as well as orbitofrontal cortex (McClure et al 2003; O'Doherty et al 2003b). In the schizophrenia literature, this approach has revealed evidence for altered striatal prediction error activity among individuals with schizophrenia using both Pavlovian and instrumental reward-learning tasks, and for both monetary and liquid rewards. An instrumental learning study with monetary reward found evidence for reduced prediction error responses among schizophrenia spectrum patients in bilateral midbrain and right ventral striatum (Murray et al 2008), and an instrumental task with liquid reward revealed reduced prediction error responses in the caudate (Gradin et al 2011). In addition, a Pavlovian task with monetary reward revealed reduced prediction error responses in the striatum in schizophrenia, coupled with enhanced responses for unsurprising reward stimuli (which are not expected to elicit prediction errors) (Morris et al 2012). In a passive paradigm that required participants to learn about the timing of a potential juice reward, Waltz and colleagues found evidence for reduced positive prediction error responses in a range of regions that included the striatum (dorsal and ventral) as well as the insula, but relatively intact negative prediction errors in these same regions. However, a study of prediction errors during the MID task revealed intact striatal prediction error activity in schizophrenia (Walter et al 2009). In all, these studies have revealed fairly consistent evidence for reduced striatal prediction errors in schizophrenia, with some suggestion

that positive prediction errors may be more affected than negative prediction errors. However, these findings are not universal, and the striatal areas demonstrating these reductions are not always consistent across studies.

In addition to these prediction error studies, two studies have used fMRI to examine reinforcement learning during more complex learning tasks. An imaging study using the Weather Prediction task found that among a sub-sample of participants defined as good learners, controls showed greater activation than individuals with schizophrenia in DLPFC and caudate, while patients showed greater activation of regions including cingulate gyrus, parahippocampal gyrus, and parietal cortex during Weather Prediction Task blocks versus perceptual-motor control task blocks (Weickert et al 2009). The authors interpret this finding as suggesting that impairments in fronto-striatal circuits may perhaps be balanced by compensatory increases in parahippocampal and parietal regions to allow preserved learning function. In a related study examining probabilistic reinforcement learning, Koch and colleagues found that controls, but not patients, had reductions in fronto-parietal activation as reward predictability increased (Koch et al 2010). Furthermore, this study also found reduced positive prediction error responses in frontal cortex, cingulate, and putamen.

There is little data on the relationship between individual differences in anhedonia/avolition and reinforcement learning performance or prediction error activity in individuals with schizophrenia. In behavioral studies, overall performance on reinforcement learning tasks has been shown to be negatively correlated with negative symptom severity (Waltz et al 2007; Yilmaz et al 2012). Further, patients with greater negative symptom severity show greater deficits in learning from positive, but not negative, feedback (Gold et al 2012). In addition, performance on Kamin blocking tasks (a behavioral indicator of prediction error signaling) has been shown to be impaired in schizophrenia, with greater impairments in performance associated with greater negative and depressive symptoms (Moran et al 2008). In neuroimaging studies, Waltz et al found that the magnitude of prediction errors in left putamen was negatively correlated with avolition scores among patients with schizophrenia during a Pavlovian conditioning task with juice rewards (Waltz et al 2009). These results suggest that there may be a relationship between negative symptoms and impairments in the ability to learn from feedback, particularly positive feedback, among individuals with schizophrenia.

In sum, the literature on reinforcement learning in schizophrenia suggests relatively intact learning on simple reinforcement learning paradigms, with evidence for overall impairment, but (perhaps) spared learning rate, on more difficult tasks. The degree to which these impairments reflect differences in striatum-influenced learning mechanisms that are more gradual and implicit vs. explicit learning mechanisms that may be more cortically mediated is an open question in the literature. Consistent with the hypothesis that some of these reinforcement learning impairments may reflect striatal mechanisms, a growing number of studies in the imaging literature suggest reduced striatal prediction error activity among this population. However, not all studies have found impaired striatal prediction error responses, and the striatal areas in which differences were detected have varied from study to study. Furthermore, at least two studies examining probabilistic reinforcement learning have also found altered activation in frontal regions, suggesting a potentially important role for cortically mediated mechanisms in these impairments. There is also evidence that the magnitude of reinforcement learning deficits and reductions in striatal prediction errors may be related to the severity of negative symptoms in this population. Several studies have suggested that reinforcement learning impairments among individuals with schizophrenia, especially those with severe negative symptoms, may be specific to learning from positive feedback, and that positive prediction errors are reduced specifically in these individuals. Together, this findings raise the possibility that impairments in learning from positive outcomes, mediated by reductions in striatal signaling of positive prediction errors, may contribute motivational deficits in schizophrenia. This possibility is examined in Chapter 4, using a probabilistic reinforcement learning paradigm allowing the separate assessment of learning from positive versus negative outcomes.

Summary

The studies reported here aim to address three possible sources of anhedonia/avolition in schizophrenia. The first, hedonics or "liking", seems to be intact in schizophrenia on the basis of self-reports of in-the-moment emotional experience and ventral striatal responses to rewarding stimuli, but may vary with individual differences in anhedonia severity. The relationships between anhedonia and self-reported hedonic experience, and its associated functional activation, are examined in Chapter 2. The second process of interest is reward prediction and "wanting". The current literature suggests that

reward anticipation in ventral striatum may be reduced in unmedicated and typically medicated patients, but spared in those taking typical antipsychotics. However, existing results may have been influenced by response requirements in addition to reward prediction processes. Chapter 3 examines reward prediction and its relationship to anhedonia/avolition using a Pavlovian paradigm with no response requirements. The third process of interest is reinforcement learning, which appears to be impaired overall for difficult tasks, but sometimes shows intact learning rates, in schizophrenia. There is disagreement in the literature as to whether striatally-mediated mechanisms of implicit learning are disrupted, with spared learning rates suggesting intact striatal learning, but reduced striatal prediction errors suggesting deficits in this process. Some data suggests that reinforcement learning impairments, particularly for patients with severe negative symptoms, may be related to specific deficits in learning from positive feedback, as well as specific reductions in positive prediction errors. This possibility is addressed in Chapter 4, using a probabilistic reinforcement learning paradigm to assess BOLD activation during learning from positive and negative feedback and its relationship to anhedonia and avolition in schizophrenia.

Chapter 2.

Anhedonia and Emotional Experience in Schizophrenia

This chapter was published in the May, 2010 issue of Biological Psychiatry. My contributions to this paper included data analysis and writing.

Reference: Dowd EC & Barch DM. Anhedonia and emotional experience in schizophrenia: neural and behavioral indicators. Biol Psychiatry (2010) vol. 67 (10) pp. 902-11.

Abstract

Background: Emotional impairments such as anhedonia are often considered key features of schizophrenia. However, self-report research suggests that emotional experience in response to affect-eliciting stimuli is intact in schizophrenia. Investigation of neural activity during emotional experience may help clarify whether symptoms of anhedonia more likely reflect alterations of in-the-moment hedonic experience or impairments in other aspects of goal-directed behavior.

Methods: 40 individuals with DSM-IV-TR schizophrenia or schizoaffective disorder and 32 healthy controls underwent fMRI while making valence and arousal ratings in response to emotional pictures, words, and faces. BOLD responses were compared between patients and controls, and were correlated with questionnaire measures of anhedonia.

Results: Patients showed some evidence of blunted valence, but not arousal, ratings in response to emotional stimuli as compared to controls. Higher anhedonia scores were associated with blunted valence ratings in both groups, and fully mediated the group differences in valence ratings. Functional activity was largely intact in patients, except for regions in right ventral striatum and left putamen, which showed reduced responses to positive stimuli. Higher anhedonia was associated with reduced activation to positive versus negative stimuli in bilateral amygdala and right ventral striatum in patients, and in bilateral caudate in controls.

Conclusions: Increased anhedonia is associated with a reduced experience of valence in both patients and controls, and group differences in experienced valence are likely driven by individual differences in anhedonia. Reduced activation of the striatum and amygdala may contribute to symptoms of anhedonia by failing to signal the salience of positive events.

Introduction

Anhedonia, or the inability to experience pleasure, is a long-established feature of schizophrenia (Bleuler 1911; Kraepelin 1971) that significantly impacts functional capacity and is resistant to treatment (Herbener et al 2005; Shoichet and Oakley 1978). Surprisingly, however, a growing body of self-report (Cohen and Minor 2008; Kring and Moran 2008) and behavioral (Heerey and Gold 2007) data suggests

that emotional experience in schizophrenia is intact. One possible explanation for this discrepancy is that in schizophrenia, clinical measures of anhedonia reflect not a deficit in consummatory pleasure, but a deficit in anticipatory pleasure or approach motivation (Burbridge and Barch 2007; Gard et al 2007; Horan et al 2006). To investigate this possibility, we asked whether neural activity during emotional experience is also intact in schizophrenia.

In studying brain responses to affect-eliciting stimuli, several structures are of particular interest. First, the *striatum* has been associated with responses to "rewarding" or pleasurable stimuli (O'Doherty 2004; Schultz 2007; Smith and Berridge 2007). Further, reduced ventral striatal activity in response to positive stimuli has been associated with anhedonia in studies with both healthy (Wacker et al 2009) and depressed (Epstein et al 2006; Keedwell et al 2005) individuals. Most commonly, the mesolimbic dopamine system and its projections to the striatum are associated with reward prediction and incentive salience (Berridge 2004; Schultz 1997; Schultz 2007), suggesting that this region is instrumental to anticipatory pleasure and approach motivation. Second, *dorsomedial prefrontal cortex* (dmPFC) and *orbitofrontal cortex* (OFC) are active during emotional experience across a wide range of emotion elicitation studies (Kober et al 2008). DmPFC may be involved in the introspective evaluation of one's feelings, while OFC may be involved in establishing the threat or reward value of a stimulus (Barrett et al 2007). Third, the *amygdala* is implicated in processing survival-salient, arousing stimuli, both negative and positive (Zald 2003). Finally, activity in the *rostral anterior cingulate cortex* (rACC) has been associated with subjective ratings of pleasantness (Rolls et al 2008a; Rolls et al 2003).

A number of studies have suggested that striatal activity during processing of positive stimuli may be altered in schizophrenia. For example, unmedicated patients have shown reduced ventral striatal activation during reward anticipation, which correlated with negative symptom severity (Juckel et al 2006b). In emotion perception studies, patients failed to modulate nucleus accumbens activation when rating pleasant versus unpleasant odors (Crespo-Facorro et al 2001), and demonstrated reduced phasic (but enhanced tonic) activity in the ventral striatum to both positive and aversive stimuli (Taylor et al 2005).

Studies of activity in dmPFC and OFC during emotional experience in schizophrenia have yielded mixed results. Given evidence of a dissociation between neural activity patterns during emotional

experience versus emotion perception (Wager 2008), we focus here on studies in which participants reported their own experienced emotion. Some of these studies found reduced activation of dmPFC and OFC during sadness in chronic and first-episode schizophrenia (Reske et al 2007; Takahashi et al 2004). However, other studies failed to find group differences in these regions in patients (Schneider et al 2007; Schneider et al 1998) and relatives (Habel et al 2004). Functional neuroimaging studies of amygdala activation in schizophrenia have also given mixed results (Aleman and Kahn 2005). Some emotional experience studies have shown reductions in amygdala activity in patients as compared to controls (Schneider et al 1998), while others found no group differences (Reske et al 2007; Schneider et al 2007). Similarly, most emotion perception studies have found reduced amygdala activity in response to emotional stimuli in patients relative to controls (Gur et al 2002b; Hempel et al 2003; Johnston et al 2005; Paradiso et al 2003; Takahashi et al 2004; Taylor et al 2002), and in paranoid vs. non-paranoid patients (Williams et al 2004). However, some studies have shown increased (Holt et al 2006; Kosaka et al 2002) or normal (Plailly et al 2006; Surguladze et al 2006) amygdala activation. The disparity in these results may reflect small sample sizes, differences in stimuli, clinical variation across samples, and the need for a low-level control condition given that neutral stimuli may elicit greater limbic activation in patients (Hall et al 2008) and their relatives (Seiferth et al 2008) than in controls. In the current study, we aimed to address these concerns by using a large sample and several types of stimuli, by conducting individual difference analyses, and by examining the pattern of responses across several emotional conditions rather than comparing emotional conditions to a neutral baseline.

This study aimed to address three questions. First, we asked whether self-reports of emotional experience are intact in patients. In keeping with previous data, we hypothesized that self-reports would be similar between patients and controls. Second, we asked whether neural activity during emotional experience was similar in patients and controls. If fMRI is sensitive to differences in emotional experience not probed by self-report measures, we would expect to see differences in functional activity in regions associated with emotional experience. Third, we asked whether there was a relationship between questionnaire measures of anhedonia and individual differences in self-reported emotion or its associated neural activity.

Methods

Participants

Participants were 40 outpatients with DSM-IV-TR schizophrenia or schizoaffective disorder and 32 healthy community controls. Controls were excluded if they had any history of, or first-order family member with, an Axis I psychotic disorder, or any current mood or anxiety disorder other than Specific Phobia. Other exclusions included: 1) DSM-IV substance abuse or dependence within six months; 2) any medical disorder that is unstable or severe, would confound the assessment of psychiatric diagnosis, or would make participation unsafe; 3) present or past head injury with neurological sequelae or causing loss of consciousness; and 4) DSM-IV mental retardation (mild or greater). The demographic and clinical characteristics of both participant groups are shown in Table 1. Groups were matched on age, parental education, gender, race, and handedness. All patients were taking antipsychotic medications, which were stable for least two weeks.

Table 1. Clinical and Demographic Characteristics

	Mean (SD)				
	CON	SCZ	t or χ^2	df	p
Age	36.25 (10.85)	36.8 (8.99)	.24	70	.82
Education	15.53 (4.29)	13.05 (2.27)	3.15	70	.002[a]
Highest Parental Education	12.76 (2.76)	13.35 (3.83)	−1.65	65	.1
Gender (% Male)	65.6	65	.003	1	.95
Race (% Caucasian)	53.1	47.5	.225	1	.64
Chapman Social Anhedonia	2.35 (2.06)	5.28 (2.17)	−5.748	69	.001[a]
Chapman Physical Anhedonia	3.45 (2.99)	7.23 (4.18)	−4.253	69	.001[a]
SANS Global Anhedonia	—	2.68 (1.21)	—	—	—
Positive Symptoms	—	1.83 (.93)	—	—	—
Negative Symptoms	—	1.81 (1.37)	—	—	—
Disorganization Symptoms	—	1.17 (.80)	—	—	—
Duration of Illness (yrs)	—	17.73 (11.25)	—	—	—
Diagnosis Subtype (%)	—		—	—	—
Schizoaffective—bipolar	—	7.5	—	—	—
Schizoaffective—depressive	—	17.5	—	—	—
Schizophrenia—undifferentiated	—	42.5	—	—	—
Schizophrenia—residual	—	15	—	—	—
Schizophrenia—paranoid	—	15	—	—	—
Schizophrenia—disorganized	—	2.5	—	—	—
Medication (% Taking)	—		—	—	—
Aripiprazole	—	22.5	—	—	—
Clozapine	—	5.0	—	—	—
Ziprasidone	—	2.5	—	—	—
Haloperidal	—	5.0	—	—	—
Iloperidone	—	2.5	—	—	—
Risperidone	—	30.0	—	—	—
Quetiapine	—	17.5	—	—	—
Trifluoperazine	—	2.5	—	—	—
Olanzapine	—	12.5	—	—	—
Average Dose (CPZ Equivalents)	—	452.20 (369.60)	—	—	—

Positive symptoms were the sum of global scores for hallucinations and delusions; negative symptoms were the sum of global scores for alogia, anhedonia, avolition, affective flattening, and attentional impairment; and disorganization symptoms were the sum of global scores for bizarre behavior, positive thought disorder, and inappropriate affect.

CON, control; SCZ, schizophrenia; SANS, scale for the assessment of negative symptoms; CPZ, chlorpromazine.

[a] $p < .05$.

Participant diagnoses were based on a Structured Clinical Interview for DSM-IV-TR (First et al 2001) and on information from medical records and corroborative sources. Clinical symptoms were rated using the Scales for the Assessment of Positive Symptoms (SAPS) (Andreasen 1983b) and Negative Symptoms (SANS) (Andreasen 1983a). We assessed anhedonia symptoms using the SANS global anhedonia score and the self-report Chapman physical and social anhedonia scales (Chapman et al 1976; Eckblad et al 1982), and handedness using the Edinburgh Index (Oldfield 1971). Patients were recruited from local psychiatric hospitals, outpatient settings, and previous studies. Patients were excluded if they had been hospitalized within the past month, or if their medications had not been stable for at least two weeks. Controls were recruited by advertising in local newspapers and posting flyers in

the St. Louis, MO, community. All subjects were paid for their participation. All clinical interviews were conducted by a Master's-level clinician who was formally trained on the SCID-IV to a between-rater reliability of 0.8. Inter-rater agreement was routinely assessed and exceeded 0.8 throughout the study using the 22 items of SCID-IV module B (psychosis and associated symptoms).

Materials and Tasks

All participants were scanned while making valence and arousal ratings of their own subjective responses to emotional pictures, words, and faces. Valence (pleasant-unpleasant) and arousal (activation-deactivation) are independent dimensions of affect (Russell 1980) that are considered vital features of emotional experience (Barrett et al 2007). Participants rated their experience of each stimulus by button press as positive, negative, or neutral during valence runs, or as highly, slightly, or not aroused during arousal runs. Stimuli consisted of 50 each emotional words, pictures, and faces, 10 in each of the following categories: negative high arousal (NHA), negative low arousal (NLA), positive high arousal (PHA), positive low arousal (PLA), and neutral (NEU). Emotional words were taken from the ANEW normed word set (Bradley and Lang 1999). Pictures were selected from the International Affective Picture System based on normed valence and arousal ratings (Lang et al 1999), and varied with respect to objects, people, locations, and actions. The faces consisted of fearful (negative), happy (positive), and neutral expressions derived from the Ekman (Ekman P 1976) and Gur (Gur et al 2002a) normed face sets. The Gur face set contains mild and extreme intensity emotions that were used for low and high arousal, respectively. For the Ekman faces, morphed images were generated between neutral and emotional expressions for each actor (LaBar et al 2003), with 50% emotion representing the low arousal condition and 100% emotion representing the high arousal condition. With the exception of the pictures, all stimuli of equal arousal were matched for valence, and all stimuli of equal valence were matched for arousal. Positive and negative pictures could not be matched for arousal because too few highly arousing images were available for use in the study. The orders of stimulus presentation and of valence versus arousal judgments were counterbalanced across participants. Participants performed the task for 6 runs; one run each of arousal and valence judgments for each stimulus type (pictures, words, faces). Stimuli were presented for 2000 msec with a jittered inter-stimulus interval varying from 1000 to 10000 msec.

Behavioral data analysis:

Valence and Arousal Ratings: We analyzed the valence and arousal ratings using repeated measures ANOVAs with group (schizophrenia, control) as a between-subjects factor and stimulus (picture, word, face) and condition (NHA, NLA, NEU, PLA, PHA) as within-subjects factors. To further characterize the pattern of responses as a function of condition, we created a priori contrasts that were sensitive to the valence and/or arousal characteristics of the stimuli. To examine the effect of valence irrespective of arousal, we used a linear contrast with weights of -1, -1, 0, 1, and 1 for NHA, NLA, NEU, PLA, and PHA, respectively (valence contrast). To examine the effect of arousal irrespective of valence, we used a quadratic contrast with weights of 2, -1, -2, -1, and 2 for NHA, NLA, NEU, PLA, and PHA, respectively (arousal contrast). To examine whether the valence ratings were influenced by both the valence and arousal characteristics of the stimuli, we also created a linear contrast in which valence was amplified by arousal, using weights of -2, -1, 0, 1, and 2, for NHA, NLA, NEU, PLA, and PHA, respectively (valenceXarousal contrast). We tested the significance of these contrasts using univariate F-tests within each group.

Individual Difference Analyses: We conducted linear regression analyses to examine the extent to which anhedonia scores predicted valence ratings within each group. To determine whether the relationship between valence ratings in response to PHA/NHA stimuli and Chapman Physical/Social Anhedonia scores differed between groups, we conducted hierarchical regression analyses with anhedonia score (physical or social) and group entered in step 1, and groupXanhedonia interaction entered in step 2. We also examined whether anhedonia scores mediated the effect of group on valence ratings. To do this, we conducted two separate multiple mediation analyses using a Sobel procedure with bootstrapping (Preacher and Hayes 2008), with PHA or NHA valence ratings as the dependent variable, group as the independent variable, and physical and social anhedonia scores as mediators. Within the patient group, we also conducted correlations between PHA/NHA valence ratings and SANS global anhedonia, and with SANS avolition, alogia, and affective flattening to examine the specificity to anhedonia.

fMRI analysis:

fMRI Acquisition and Image Analysis: All scans were performed on a 3T Siemens Allegra head-only system. We acquired structural images using a sagittal T1-weighted MP-RAGE sequence [TE = 2.9 ms, TR = 6.6 ms, flip angle = 8˚, acquisition matrix = 96 x 128, 80 slices, 2 x 2.67 x 2 mm voxels]. To facilitate registration of the T1 and functional scans, we also acquired a T2 image in the same space as the functional scans [TE = 96 ms, TR = 5 s, 189 x 256 acquisition matrix, 48 slices, 1.02 x 1 x 3 mm voxels]. The functional images were collected in runs using an asymmetric spin-echo echo-planar sequence sensitive to blood oxygenation level-dependent (BOLD) contrast (T2*) [TR = 3000 ms, TE = 25 ms, FOV = 205 mm, flip = 90°, 40 axial slices, 3.2 mm^3 isotropic voxels]. Stimuli were presented using PsyScope on a G3 Macintosh computer, with each trial onset triggered directly by a pulse from the scanner. A fiber-optic button box interfaced with PsyScope was used to record participants' responses.

The fMRI data was preprocessed and analyzed using in-house software. The functional images were first normalized across runs by scaling whole-brain signal intensity to a fixed value and removing the linear slope on a voxel-by-voxel basis to counteract effects of drift (Bandettini et al 1993). The MR data was then aligned to correct for head motion using rigid-body rotation and translation correction algorithms (Snyder 1996; Woods et al 1992). Frame-to-frame movement and signal-to-noise ratio were compared between groups, and subjects showing excessive movement or poor signal quality were excluded. The structural and functional scans were registered to Talairach space (Talairach and Tournoux 1988) using a 12 parameter linear (affine) transformation (Woods et al 1992), and smoothed with a 6 mm FWHM Gaussian filter.

Event-related analyses were used to obtain estimates of activation during the five conditions (NHA, NLA, PHA, PLA, and NEU) for each stimulus type (picture, word, face). For each participant, a general linear model (GLM) (Friston et al 1995) was used to estimate a hemodynamic response function for each trial type. The GLM included regressors for linear trends within runs and baseline shifts between runs. An assumed hemodynamic response shape (Boynton function) was used to generate magnitude estimates for each event type, and these magnitude estimates were used in all further statistical analyses. Functional activation was analyzed using the valence, arousal, and valenceXarousal contrasts in both region-of-interest (ROI) and whole-brain analyses. We examined voxelwise t-tests at the group

level within predefined ROI masks including the amygdala, striatum, dmPFC, OFC, and rACC. The amygdala and basal ganglia ROIs were derived from manually outlined anatomical templates (Mamah et al 2007; Wang et al 2008) that were projected into Talairach space, and the dmPFC, rACC, and OFC ROIs were 15 mm diameter spherical ROIs centered on the coordinates reported in (Kober et al 2008). The ROI mask is shown in supplemental figure S1. Both the whole-brain and ROI analyses were corrected for multiple comparisons using combined p-value/cluster size thresholds, determined using Monte Carlo simulations to provide an overall false-positive rate of 0.05 (Forman et al 1995; McAvoy et al 2001). These thresholds were $p<.01$ and 14 voxels for ROI analyses, and $p<.003$ and 30 voxels for whole-brain analyses. To identify regions whose activation patterns were consistent with the valence and arousal patterns of interest, we first conducted one-sample t-tests for each contrast on both groups combined. To identify regions showing group differences in activation, we also performed group t-tests on each contrast. In both analyses, significant regions were followed up with simple effects tests to determine the activation pattern within each group separately. To eliminate redundancy, regions identified by both the valence and valence X arousal contrasts were treated as follows. To determine which contrast better described the activation pattern in such regions, we performed post-hoc pairwise comparisons between the activation magnitudes for each condition. If the region showed both a valence effect (positive <> negative) and an arousal effect (NHA <> NLA or PHA <> PLA), it was evaluated with the valence X arousal contrast and removed from the valence contrast map. If it showed only a valence effect, it was removed from the valence X arousal map and evaluated with the valence contrast. To examine individual differences in functional activity, we also conducted voxelwise correlation analyses between contrast scores and anhedonia scores. Correlations were conducted within each group separately, and correlation coefficients were compared between groups using Fisher r-to-z transformations.

In addition to the analysis using canonical hemodynamic response shapes, a separate analysis was conducted that estimated response magnitudes for each MR frame separately, as in (Miezin et al 2000). Comparable results were obtained using this analysis, and example timecourses are shown in Figure S5.

Results

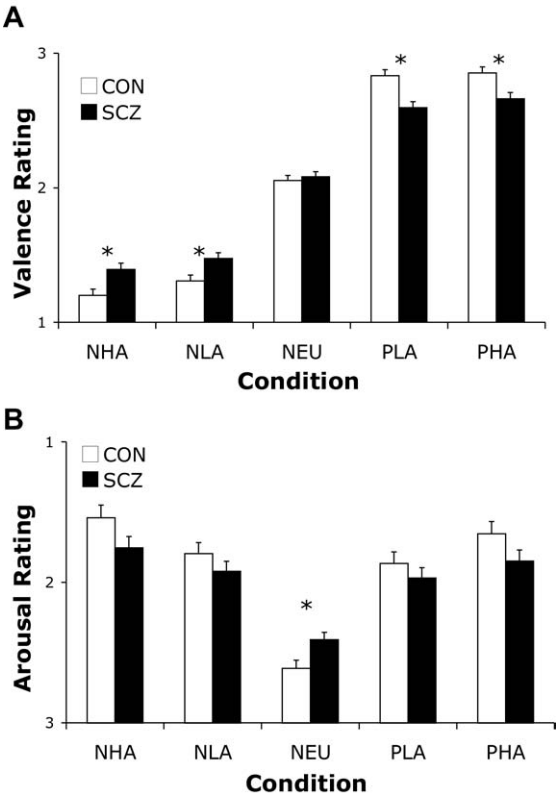
Behavioral Results:

Anhedonia Scores: Overall, individuals with schizophrenia had higher anhedonia scores than controls on both Chapman scales (Table 1).

Valence and Arousal Ratings: For valence (Figure 1a), there was a significant main effect of condition ($F(4, 280) = 659.85$, $p<.001$) and significant groupXcondition ($F(4, 280) = 13.49$, $p<.001$) and stimulusXcondition ($F(8, 560) = 12.70$, $p<.001$) interactions. Simple effects tests revealed significant effects of condition for controls ($F(4, 280) = 386.85$, $p<.001$) and patients ($F(4, 280) = 273.95$, $p<.001$). However, comparisons within each condition revealed that patients' responses to negative stimuli were less negative ($F(1,70) = 9.62$, $p <.004$ for NHA; $F(1,70) = 7.87$, $p<.007$ for NLA), and to positive stimuli were less positive ($F(1,70) = 15.77$, $p<.001$ for PLA, $F(1,70) = 9.75$, $p<.004$ for PHA), than controls. Both the valence and valenceXarousal contrasts were significant for both groups, with similar effect sizes (valence contrast: $F(1,70) = 521.50$, $p<.001$, $\eta_p^2 = 0.882$ for controls, $F(1,70) = 368.23$, $p<.001$, $\eta_p^2 = 0.840$ for patients; valenceXarousal contrast: $F(1,70) = 509.33$, $p<.001$, $\eta_p^2 = 0.879$ for controls; $F(1, 70) = 364.80$, $p<.001$, $\eta_p^2 = 0.839$ for patients). Because stimulus type (picture, word, face) did not interact with group in any of our analyses (behavioral or fMRI), stimulus effects are not discussed further here (but see *Supplemental Materials*).

For the arousal ratings (Figure 1b), there were significant main effects of stimulus ($F(2, 140) = 23.41$, $p<.001$) and condition ($F(4, 280) = 55.96$, $p<.001$), and significant stimulusXcondition ($F(8, 560) = 5.03$, $p<.001$) and groupXcondition $F(4, 280) = 3.15$, $p<.009$) interactions. Simple effects tests revealed significant effects of condition within each group ($F(4,280) = 39.35$, $p<.001$ for controls, $F(4,280) = 17.72$, $p<.001$ for patients). Further, group comparisons within each condition revealed a significant group difference only for the NEU condition: compared to controls, patients showed higher arousal in response to neutral stimuli ($F(1,70) = 6.87$, $p< 0.02$). The arousal contrast was significant for both groups ($F(1, 70) = 151.45$, $p<.001$, $\eta_p^2 = 0.654$ for controls; $F(1, 70) = 68.76$, $p<.001$, $\eta_p^2 = 0.496$ for patients). Taken together, these results indicate that patients showed blunted valence ratings in response to emotional stimuli. However, the patterns of both valence and arousal ratings as a function of emotional condition were similar between groups.

Figure 1. Valence and arousal ratings as a function of condition in individuals with schizophrenia (SCZ) and control subjects (CON). (A) Average valence ratings (1 = negative, 2 = neutral, 3 = positive) collapsed across stimulus type (pictures, words, faces) for each emotional condition: negative high arousal (NHA), negative low arousal (NLA), neutral (NEU), positive low arousal (PLA), and positive high arousal (PHA). (B) Average arousal ratings (1 = highly aroused, 2 = slightly aroused, 3 = not aroused) collapsed across stimulus type (pictures, words, faces) for each emotional condition. *p < .05. Error bars represent standard error.



Individual Difference Analyses: We conducted hierarchical regression analyses using Chapman anhedonia scores to predict valence ratings to PHA and NHA stimuli in patients and controls (Table 2). In all of these analyses, anhedonia score and group accounted for a significant portion of the variance in the valence ratings, and adding a groupXanhedonia interaction term failed to account for significantly more variance. As expected, higher physical and social anhedonia scores were associated with less-positive responses to PHA stimuli and less-negative responses to NHA stimuli in both groups (Figure 2). Similarly, within the patient group, SANS anhedonia correlated negatively with PHA valence ratings ($r = -.37$, $p<.03$), though it failed to correlate with NHA valence ratings ($p>.16$). Together, these results suggest that within both patients and controls, higher levels of anhedonia are associated with less-valenced experiences of emotional stimuli.

Multiple mediation analyses revealed that for both PHA and NHA valence ratings, the effect of group was fully mediated by physical and social anhedonia scores (effect of group on valence rating, controlling for physical and social anhedonia: t(69) = -1.2, p>.24 for PHA, t(69) = 1.0, p>.31 for NHA). The total mediated effect was significant in both models (95% CI = -0.27, -0.01 for PHA; 0.02, 0.25 for NHA). For PHA ratings, only physical anhedonia was significant as a specific mediator (95% CI = -0.16, -0.003), and for NHA ratings, neither specific mediator was significant alone.

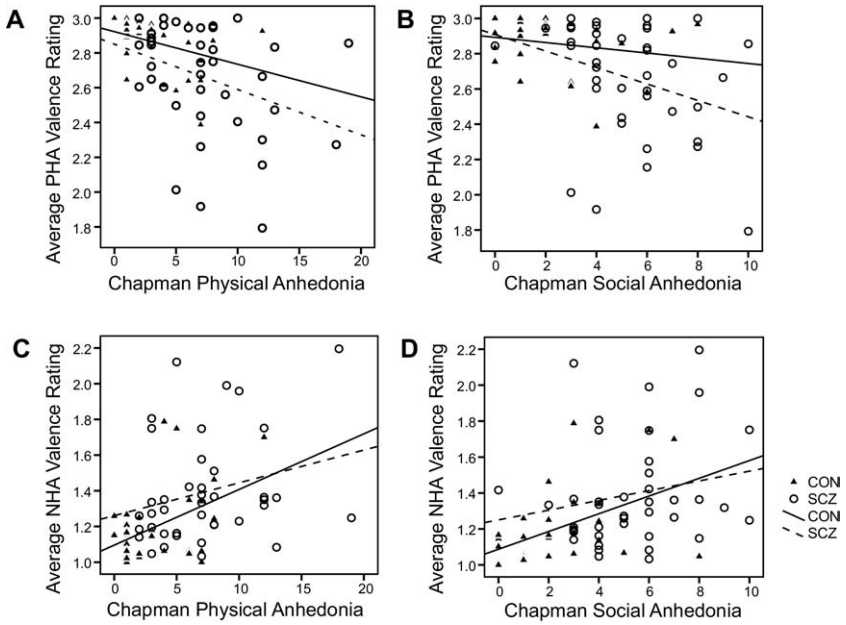
To evaluate the specificity of these results to anhedonia, we correlated PHA and NHA valence ratings with SANS global avolition, alogia, and affective flattening in patients, and found that avolition also correlated negatively with PHA valence ratings (r = -.34, p<.04) and positively with NHA valence ratings (r = .36, p<.03). Aside from a trend-level correlation between alogia and NHA ratings (r = .30, p<.07), alogia and affective flattening failed to correlate with either measure (p>.16). Therefore, a reduced experience of positive and negative emotion appears to be related to symptoms of anhedonia and amotivation, but not to other negative emotional symptoms.

Table 2. Hierarchical Regression Analyses with Group and Anhedonia Scores to Predict Valence Ratings

Dependent Variable	Predictors	R^2	R^2 Change	β
PHA Valence Rating	Step 1	.229[a]		
	Group			−.191
	Physical anhedonia			−.360[b]
	Step 2	.231[b]	.002	
	Group × physical anhedonia			−.127
PHA Valence Rating	Step 1	.192[b]		
	Group			−.177
	Social Anhedonia			−.313[d]
	Step 2	.206[b]	.014	
	Group × social anhedonia			−.357
NHA Valence Rating	Step 1	.199[b]		
	Group			.191
	Physical anhedonia			.326[c]
	Step 2	.205[b]	.006	
	Group × physical anhedonia			−.219
NHA Valence Rating	Step 1	.190[b]		
	Group			.151
	Social anhedonia			.332[d]
	Step 2	.197[c]	.007	
	Group × social anhedonia			−.246

PHA, positive high arousal; NHA, negative high arousal.
[a] $p < .001$.
[b] $p < .005$.
[c] $p < .01$.
[d] $p < .05$.

Figure 2. Scatterplots of (A) average valence ratings to PHA stimuli as a function of Chapman physical anhedo- nia, (B) average valence ratings to PHA stimuli as a func- tion of Chapman social anhedonia, (C) average valence ratings to NHA stimuli as a function of Chapman physical anhedonia, and (D) average valence ratings to NHA stim- uli as a function of Chapman social anhedonia, in individ-uals with SCZ (circles) and CON (triangles). Abbreviations as in Figure 1.



fMRI Results:

Movement and Signal-to-Noise Ratio: Three patients and three controls were excluded from analysis for excessive movement and/or low signal-to-noise ratio (SNR). Mean incremental (frame-to-frame) movement was computed for each run for each subject and used to compare movement between groups. For the final sample of 40 patients and 32 controls, mean incremental movement and group *t*-test results for each translation axis (x, y, z) and rotation axis (pitch, roll, yaw) are summarized in Table S1. The groups differed significantly on movement in the y axis only. SNR was computed by determining the ratio of the mean signal intensity to its standard deviation for each frame within a run, from which the mean, median, and maximum SNR values were determined for each run for each participant. The mean SNR values across runs were then calculated and compared between groups (Table S2). Despite the group difference in y-axis movement, SNR did not differ significantly between groups. Together, these results indicate that the groups were well matched for signal quality and that poor signal quality in patients is unlikely to contribute to the group results reported here.

ROI Analyses: One-sample t-tests: As shown in Figure 3, one-sample t-tests identified several regions within the ROIs with activity patterns significant for the valence, arousal, and valenceXarousal contrasts. These regions and their activation patterns are detailed in Table 3.

Table 3. Results of ROI Analyses of Valence, Arousal, and Valence × Arousal Contrasts in the Total Sample

Brain Region	Talairach Coordinates	BA	Number Voxels	*Z*	Activation Pattern		Regional Analyses				
					Valence	Arousal	Main Effect of Group	Group × Condition Interaction	Group Difference in Contrast	Within CON Contrast	Within SCZ Contrast
Valence Contrast											
R OFC	41, 24, −4	47	27	−3.98	neu > neg > pos	—	[b]	NS	NS	[b]	[b]
rACC	1, 35, 3	32	17	4.61	deactivation: neu = neg > pos	—	NS	NS	NS	[a]	[c]
Arousal Contrast											
R amygdala	26, −14, −8	—	24	3.39	—	ha > la = neu	NS	NS	NS	[b]	[d]
dmPFC	−1, 51, 23	9	79	3.68	—	ha > la = neu	NS	NS	NS	NS	[b]
Valence × Arousal Contrast											
R amygdala	28, −10, −15	—	21	−3.41	neg > neu = pos	nha > nla = neu	NS	NS	NS	[d]	[c]

ROI, region of interest; BA, Brodmann area; CON, control; SCZ, schizophrenia; R, Right; OFC, orbitofrontal cortex; neu, neutral; neg, negative; pos, positive; rACC, rostral anterior cingulate cortex; ha, high arousal; la, low arousal; dmPFC, dorsomedial prefrontal cortex; nha, negative high arousal; nla, negative low arousal.

[a] $p < .001$.

[b] $p < .005$.

[c] $p < .01$.

[d] $p < .05$.

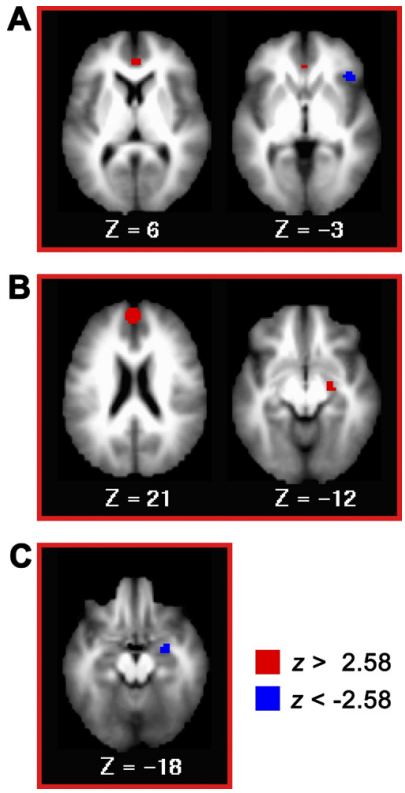


Figure 3. Results of region of interest analyses of valence, arousal, and valence X arousal contrasts in the total sample (both patients and control subjects). Regions are described in Table 3. Images are shown in neurologi- cal orientation. (A) Regions showing significant activation in the valence contrast. Positive z scores (red) indicate greater activation to positively valenced stimuli than to negatively valenced stimuli; negative z scores (blue) indicate greater activation to negatively valenced stimuli than to positively valenced stimuli. (B) Regions showing significant activation in the arousal contrast. Positive z scores (red) indicate greater activation to high arousal stimuli relative to neutral and low-arousal stimuli. (C) Regions showing significant activation in the valence X arousal contrast. Negative z scores indicate greater activation to negative high arousal stimuli than to positive and/or low-arousal stimuli.

Group t-tests: One region in right ventral striatum demonstrated a significant group difference in the valence contrast, and one in left putamen showed a group difference in the valenceXarousal contrast (Table 4 and Figure 4). As shown in Figure 4, in both of these regions, patients showed altered activation as compared to controls for the positive conditions. Post-hoc tests revealed that in left putamen, PHA activation was significantly reduced in patients ($F(1,70)=5.05$, $p<.04$). In right ventral striatum, there were significant group differences in both PHA ($F(1.70)=5.58$, $p<.03$) and PLA ($F(1.70)=6.14$, $p<.03$). These regions showed significant activation in controls, but no significant change from baseline in patients.

Figure 4. Results of region of interest analyses of group t tests between patients and control subjects for the valence and valence X arousal contrasts. Regions are described in Table 4. Images are shown in neurological orientation. (A) Right ventral striatal region demonstrating a group difference in the valence contrast. Activation did not differ significantly in the NHA, NLA, or NEU conditions but was significantly lower in patients than in control subjects for the PLA and PHA conditions. (B) Left putamen region demonstrating a group difference in the valence X arousal contrast. Activation did not differ significantly in the negative, neutral, or PLA conditions but was significantly lower in patients than in control subjects in the PHA condition. *p < .05. Error bars represent standard error. Abbreviations as in Figure 1.

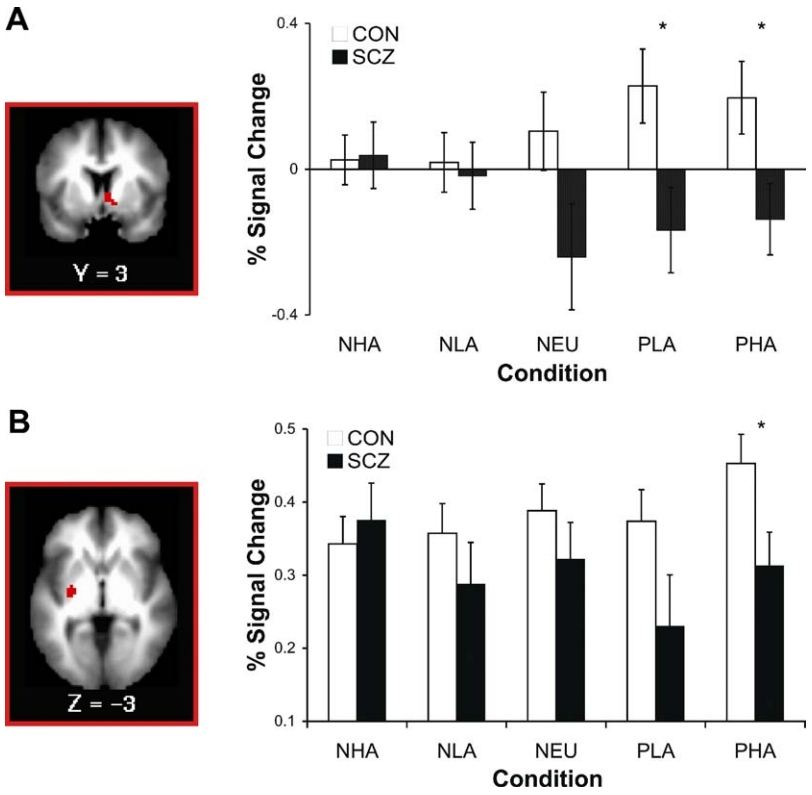


Table 4. Results of ROI Analyses of Group *t* Tests Between Patients and Control Subjects for the Valence and Valence × Arousal Contrasts

Brain Region	Talairach Coordinates	Number Voxels	*Z*
valence contrast			
R ventral striatum	7, 3, −4	14	2.80
Valence × Arousal Contrast			
L putamen	−28, −14, −1	17	3.44

ROI, region of interest; R, Right; L, Left.

Whole Brain Analyses:

As shown in Table S5 and Figure 5, a number of regions were identified by the valence, arousal, and valenceXarousal contrasts in whole-brain one-sample t-tests. In the group t-tests, however, we did not find a single region that showed a significant group difference in any of the contrasts. To further

examine whether the activity patterns were similar between groups, we conducted follow-up group analyses on each region identified in the one-sample t-tests. As shown in Table S5, overall activity differed between patients and controls in a number of these regions. However, in every region: (1) both patients and controls showed significant within-group effects of the relevant contrast, (2) there were no significant group differences in the magnitude of the contrast; and in all but two regions, (3) the pattern as a function of emotional condition was the same for both patients and controls. Thus, outside of the striatum, patients and controls demonstrated similar neural responses to both valence and arousal.

Figure 5. Results of whole-brain analyses of valence, arousal, and valence X arousal contrasts in the total sample (both patients and control subjects). Regions are described in Table S5 in Supplement 1. Images are shown in neurological orientation. (A) Regions showing significant activation in the valence contrast. Positive z scores (red) indicate greater activation to positively valenced stimuli than to negatively valenced stimuli; negative z scores (blue) indicate greater activation to negatively valenced stimuli than to positively valenced stimuli. (B) Regions showing significant activation in the arousal contrast. Positive z scores (red) indicate greater activation to high arousal stimuli relative to neutral and low-arousal stimuli; negative z scores (blue) indicate greater activation to neutral and low-arousal stimuli than to high-arousal stimuli. (C) Regions showing significant activation in the valence X arousal contrast. Negative z scores indicate greater activation to negative high arousal stimuli than to positive and/or low-arousal stimuli.

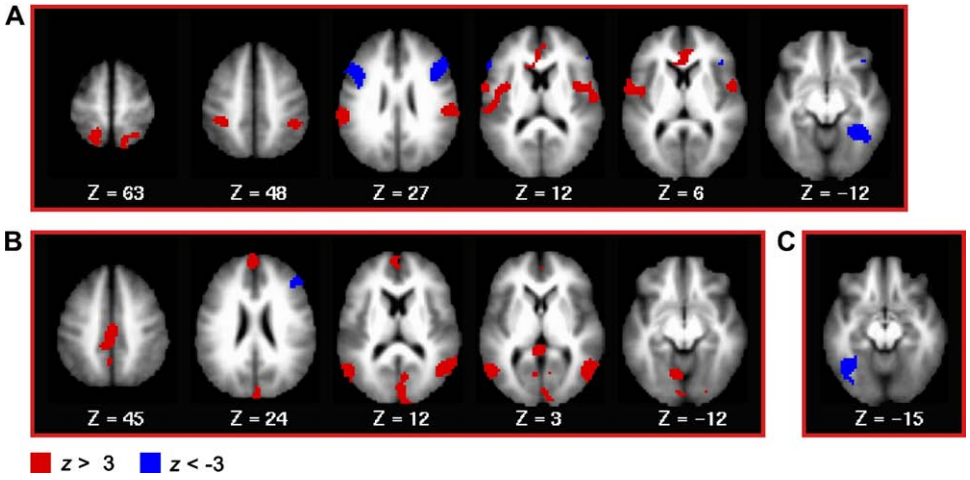


Table S5: Results of whole-brain analyses of valence, arousal, and valenceXarousal contrasts in the total sample

Brain Region	BA	Talairach Coordinates	# Voxels	*Z*	Activation Pattern		Regional Analyses				
					Valence	Arousal	Main Effect of Group	Group X Condition Interaction	Group Difference in Contrast	Within CON Contrast	Within SCZ Contrast
Valence Contrast:											
Midbrain	-	1, -27, -3	45	-4.50	neg > neu=pos	-	***	*	NS	****	*
R Fusiform Gyrus	37	39, -50, -13	144	-3.66	neg=neu > pos	-	NS	NS	NS	****	****
R Orbital Frontal Cortex	47	45, 28, -4	82	-3.25	neg=neu > pos	-	***	*	NS	***	****
R Inferior Frontal Gyrus	9	45, 15, 29	242	-3.98	neg=neu > pos	-	***	NS	NS	****	****
L Inferior Frontal Gyrus	9	-47, 13, 26	163	-2.96	neg=neu > pos	-	***	NS	NS	***	****
L Postcentral Gyrus	40	-54, -24, 23	516	4.02	pos > neu=neg	-	NS	NS	NS	****	****
R Central Operculum	43	51, -5, 11	130	3.56	pos=neu > neg	-	**	NS	NS	****	***
R Supramarginal Gyrus	40	59, -29, 30	167	3.61	pos=neu > neg	-	*	NS	NS	****	****
R Supramarginal Gyrus	40	45, -43, 49	69	2.69	pos=neu > neg	-	NS	NS	NS	**	***
R Superior Parietal Lobule	7	16, -60, 63	88	2.46	pos=neu > neg	-	**	NS	NS	*	****
L Superior Parietal Lobule	7	-18, -55, 64	70	2.89	pos=neu > neg	-	*	NS	NS	***	***
Rostral Anterior Cingulate	24	-1, 31, 6	112	3.98	deactivation: neg > pos	-	NS	NS	NS	****	**
Arousal Contrast:											
R Middle Occipital Gyrus	37	53, -64, 7	210	5.26	-	ha > la > neu	NS	NS	NS	****	****
R Cerebellar Declive	-	3, -77, -14	191	4.65	-	ha > neu=la	NS	NS	NS	*	****
R Cuneus	18	5, -90, 10	205	4.33	-	ha > neu=la	NS	NS	NS	***	***
L Posterior Cingulate	29	-2, -43, 2	60	3.87	-	ha > neu=la	*	NS	NS	***	*
L Middle Occipital Gyrus	37	-54, -66, 6	114	4.45	-	ha > neu=la	NS	NS	NS	****	*
R Middle Frontal Gyrus	9	46, 31, 27	49	-4.21	-	neu > la > ha	**	NS	NS	***	***
Dorsomedial Prefrontal Cortex	9	-1, 54, 21	140	4.20	-	deactivation: neu > la; activation: ha	NS	NS	NS	*	****
L Paracentral Lobule	31	-2, -28, 45	133	4.86	-	deactivation: neu=la; activation: ha	NS	NS	NS	****	***
L Precuneus	7	-1, -55, 49	30	4.25	-	deactivation: neu=la > ha	NS	NS	NS	***	***
ValencexArousal Contrast:											
L Fusiform Gyrus	37	-39, -62, -15	123	-4.78	neg=neu > pos	nha > nla=neu	NS	NS	NS	*	****

* p<.05, ** p<.01, *** p<.005, **** p<.001; CON = control; SCZ = schizophrenia; R = Right; L = Left; pos = positive; neg = negative; neu = neutral; ha = high arousal; la = low arousal; NS = not significant; BA = Brodmann Area

Individual difference analyses: We first conducted correlations between anhedonia scores and average activation contrast scores within the regions showing group differences in the contrasts. This analysis revealed a negative correlation between physical anhedonia and valenceXarousal contrast score in the right ventral striatum in patients (r = -.36, p<.04), indicating that patients with higher anhedonia scores showed less activation in this region in response to positive stimuli as compared to neutral and negative stimuli. This correlation was not significant in controls (r = -.17, p>.34), though the group difference in correlation coefficients was not significant (p>.77). We next conducted voxelwise ROI analyses (Table 5), in which physical anhedonia correlated negatively with the valence contrast in left amygdala, and with the valenceXarousal contrast in right amygdala, in patients. In controls, social anhedonia correlated negatively with the valence contrast in bilateral caudate. Comparison of correlation coefficients between groups revealed a significant difference in the right caudate (p<.02) and a trend level difference in left caudate (p <.08), but no difference in either amygdala region (p>.70).

Table 5. ROI Results of Correlation Analyses Between Anhedonia Scores and fMRI Valence and Valence × Arousal Contrast Scores

Contrast	Anhedonia Measure	Talairach Coordinates	Region Name	Number Voxels	*r*	*Z*
Control						
Valence	Chapman social anhedonia	8, −1, 14	R caudate	14	−.550	−3.26
		−13, −2, 18	L caudate	28	−.539	−3.18
Schizophrenia						
Valence	Chapman physical anhedonia	−14, −9, −15	L amygdala	17	−.446	−2.89
Valence × Arousal	Chapman physical anhedonia	17, −14, −11	R amygdala	15	−.413	−2.58

ROI, region of interest; fMRI, functional magnetic resonance imaging; R, Right; L, Left.

Medication Effects on fMRI Data: In a sample restricted to the 24 patients who were taking only atypical antipsychotics (excluding risperidone), group *t*-tests within the two regions that showed group differences in the full sample remained significant. In right ventral striatum (7, 3, -4; Figure S3), there were significant group differences in the valence contrast (t(54) = 2.25, p = .03). The group differences within the positive and neutral conditions dropped to trend-level significance (F(1,54) = 3.17, p = .08 for NEU; F(1,54) = 3.25, p = .08 for PLA; F(1,54) = 2.75, p = .10 for PHA), perhaps reflecting reduced power after exclusion of the 16 patients taking typical antipsychotics or risperidone. In left putamen (-28, -14, -1; Figure S4), there was a significant group difference in the valence X arousal contrast (t(54) = 3.19, p = .003). The group difference in PHA dropped to trend level (F(1,54) = 3.67, p = .06). The correlations between physical anhedonia and activation within the right ventral striatum and bilateral amygdala remained significant in the subset of patients taking only atypical antipsychotics. In ventral striatum,

physical anhedonia correlated significantly with the valence X arousal contrast (r = -.454, p = .03). Physical anhedonia correlated with the valence contrast in left amygdala (r = -.542, p = .006), and with the valence X arousal contrast in right amygdala (r = -.620, p = .001).

When we conducted correlations between antipsychotic dosage in chlorpromazine equivalents and activation in each contrast within the left putamen and right ventral striatal regions, none of the correlations reached significance. We also conducted a voxelwise correlation within our full set of ROIs, and found only one region in the right dorsal caudate that showed a significant correlation in the arousal contrast (coordinates: 16, -14, 21; 23 voxels, r = .514, p < .001.) This region did not overlap with the regions that showed group differences in activity in the valence and valence X arousal contrasts.

Discussion

Behavioral Measures of Emotional Experience

In agreement with most clinical data, we found that individuals with schizophrenia self-reported more anhedonia than controls. Behaviorally, there was a groupXcondition interaction in the valence ratings, and post-hoc tests revealed that patients rated their experience of the valenced stimuli as less valenced than controls. This finding is at odds with the majority of studies, which have shown intact responses to emotional stimuli. However, while the arousal ratings also showed a groupXcondition interaction, the only post-hoc group difference was heightened arousal ratings in response to neutral stimuli in patients. This finding suggests that patients' experience of arousal in response to emotional stimuli is intact, in agreement with previous literature. Furthermore, patients clearly showed modulation of both valence and arousal ratings as a function of the emotional content of the stimuli: when we conducted contrast analyses sensitive to valence, arousal, and valenceXarousal interaction, the relevant contrasts were significant within both groups, with similar effect sizes. Overall, while these findings suggest that the range of experienced emotion may be narrowed in patients, they also show that evoked arousal is relatively intact, and that affective stimuli modulate emotional experience in similar ways in patients and controls.

Individual difference analyses revealed that higher anhedonia was associated with blunted responses to emotional stimuli within both patients and controls. Furthermore, the group differences in

valence ratings were fully mediated by anhedonia scores. Together, these results indicate that the level of anhedonia, rather than simply the diagnosis of schizophrenia, may underlie the blunted responses to emotional stimuli seen in patients. This finding highlights the importance of including sufficiently powered individual difference analyses in future work.

fMRI Measures of Emotion Processing

fMRI analysis revealed that brain activity is largely intact during emotional experience in schizophrenia. On whole-brain analysis, we did not find any regions that showed group differences in any contrast, suggesting similar patterns of neural activity in patients and controls. On ROI analysis, however, right ventral striatum and left putamen showed reduced activation to positive stimuli in patients as compared to controls. Given past research showing that striatal activation is associated with the anticipation (O'Doherty et al 2002) and receipt (Rolls et al 2008a) of pleasurable stimuli, this finding may represent a failure to respond to positive experiences that contributes to an inability to anticipate or want such experiences in the future (Berridge and Kringelbach 2008).

In support of this interpretation, reduced activation to positive versus negative stimuli in the same ventral striatal region was also associated with higher physical anhedonia in patients. This finding suggests that the group differences in activation seen in this region may be driven by individual differences in anhedonia. Similarly, bilateral amygdala activation to positive versus negative stimuli was reduced in patients who where higher in physical anhedonia. Within controls, greater social anhedonia was associated with decreased bilateral caudate activation in response to positive relative to negative stimuli. Because the amygdala (Zald 2003) and striatum (Zink et al 2004) are thought to be involved in salience attribution, these results may indicate that these regions fail to mark positive events as salient in anhedonic individuals, leading to a blunted experience of emotion and a reduced ability to seek out similar events in the future.

Given that the ventral striatum is typically associated with reward processing, it is interesting to speculate on how these results relate to findings of reduced ventral striatal activation during reward anticipation in schizophrenia (Juckel et al 2006b). Notably, the reduced ventral striatal activation to positive stimuli seen here in patients may represent a deficit in motivational or reward-prediction processes, rather than in hedonic processes per se. During learning, dopaminergic neurons initially fire to

unexpected positive stimuli, shifting over time to fire to cues that predict these rewards (Schultz 1997). Thus, the deficient right ventral striatal activation reported here could reflect a failure of this initial dopaminergic firing to unpredicted positive stimuli, potentially impairing reward prediction/incentive salience and leading to reduced anticipatory activation. Importantly, this impairment in predictive or motivational processes may be independent of the hedonic response to the reward, allowing a normal experience of "liking" combined with reduced "wanting". This is consistent with the view that consummatory pleasure is intact in schizophrenia while anticipatory pleasure is impaired (Gard et al 2007).

Given the finding of group differences in striatal activity, a major limitation of this study is that all patients were taking medications that block dopamine receptors, potentially altering striatal function. However, the majority of patients were taking atypical antipsychotics, which have a lesser effect on striatal activity during reward processing than typical antipsychotics (Juckel et al 2006a). Further, when we removed from analysis all patients taking typical antipsychotics or risperidone (which are pharmacologically similar), the group differences and correlations remained significant. In addition, functional activation did not correlate with antipsychotic dose within the regions showing group differences. While the possibility of medication effects cannot be ruled out without examination of unmedicated patients, we feel that these results provide reasonable evidence that the findings reported here were not driven by medications.

In summary, this study makes several important contributions to the literature on emotional experience and its related brain activity in schizophrenia. First, while patients showed blunted responses to emotional stimuli as compared to controls, these group differences in ratings were clearly mediated by the level of anhedonia displayed by the participants. Second, the pattern of brain activity in response to emotional stimuli was largely intact, with the exception of two striatal regions that showed reduced responses to positive stimuli. Third, blunted activation to positive vs. negative stimuli correlated with anhedonia in the amygdala and right ventral striatum in patients, and in the caudate in controls, suggesting that failure to mark stimuli as salient or rewarding may contribute to symptoms of anhedonia. Clinically, these results highlight the importance of individual differences, suggesting that optimal treatment strategies are best tailored to the individual symptomatology of the patient. Future work

examining the relationship between reduced neural responses to positive stimuli and deficits in motivated behavior, using paradigms that probe for reward anticipation and reinforcement learning in anhedonic individuals, may shed additional light on the questions raised here.

Supplemental Information

Example Task Instructions:

Picture Rating Instructions for Valence:

"In this part of the study, I am going to show you some pictures. Some of the pictures will be positive, some will be negative, and some will be neutral. I want you to decide how you personally feel about the pictures, and then to make a response based on how the picture makes you feel. What I want you to do is to press the button with your pointer finger every time you decide a picture is positive. Press the button with your middle finger every time you decide a picture is neutral. Press the button with your ring finger every time you decide a picture is negative."

Picture Rating Instructions for Arousal:

"In this part of the study, I am going to show you some pictures. Some of the pictures will be highly arousing, some will be only a little arousing, and some will not be arousing at all. A picture is highly arousing if it causes you to feel strong emotion about it. That is, you really have strong feelings about what the picture means. A picture is only a little arousing when you feel some emotion about it, but it is not that strong. A picture is not arousing at all when you do not feel any emotion toward the word. Both positive and negative pictures can be highly arousing or only a little arousing. I want you to decide how you personally feel about the pictures, and then to make a response based on how the picture makes you feel. What I want you to do is to press the button with your pointer finger every time you decide a picture is highly arousing. Press the button with your middle finger every time you decide a picture is not arousing at all. Press the button with your ring finger every time you decide a picture is only a little arousing."

Regions of Interest

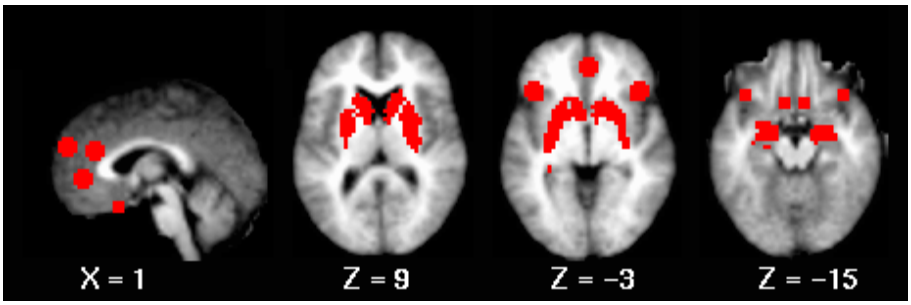


Figure S1. Regions used in ROI analysis

Movement analyses

Table S1. Incremental movement

	Control		Schizophrenia		
	Mean	SE	Mean	SE	*p*
x	0.022	0.010	0.046	0.009	0.092
y	0.037	0.005	0.055	0.004	0.010*
z	0.051	0.010	0.077	0.009	0.052
pitch	0.052	0.009	0.072	0.008	0.097
roll	0.025	0.007	0.039	0.006	0.131
yaw	0.021	0.006	0.036	0.006	0.080

SE, standard error
* represents p < .05

Table S2. Signal-to-noise ratio

	Control		Schizophrenia		
	Mean	SE	Mean	SE	*p*
Mean SNR	340.286	37.116	326.077	33.197	0.776
Median SNR	327.345	19.875	318.206	17.777	0.733
Max SNR	720.273	277.86	789.291	248.53	0.854

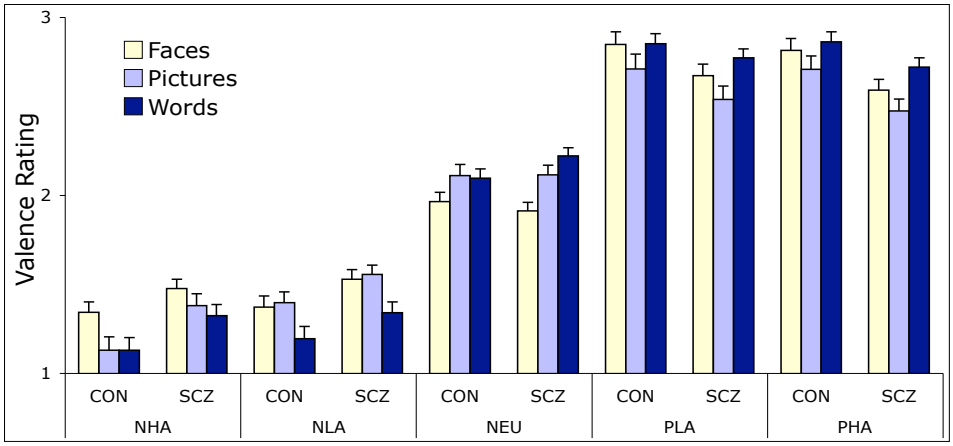
SE, standard error; SNR, signal-to-noise ratio

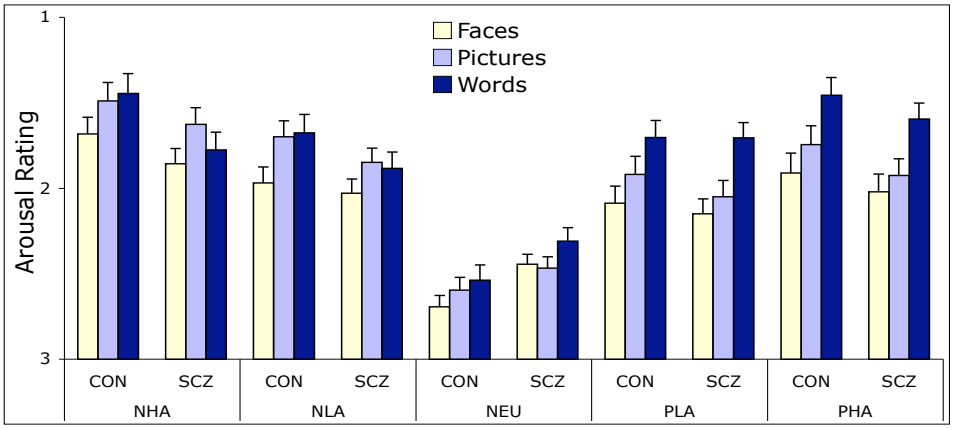
Effects of Stimulus Type

Behavioral: Valence and arousal ratings were evaluated with separate repeated measures ANOVAs with stimulus (picture, word, face) and condition (NHA, NLA, NEU, PLA, PHA) as within-subjects

factors and group (schizophrenia, control) as a between-subjects factor. For valence (Figure S2A), there was a significant stimulus-by-condition interaction ($F(8,560) = 12.70$, $p < .001$). There was no main effect of stimulus ($F(2,140) = 0.97$, $p > .37$), interaction of stimulus with group ($F(2,140) = 0.77$, $p > .46$), or three way interaction between stimulus, condition and group ($F(8,560) = 0.52$, $p > .74$). Simple effects tests to follow up on the stimulus-by-condition interaction revealed significant effects of condition within each stimulus (faces: $F(5,350) = 2577$, $p < .001$; pictures: $F(5,350) = 1930$, $p < .001$; words: $F(5,350) = 2739$, $p < .001$), and significant effects of stimulus within each condition (NHA: $F(3,210) = 811.27$, $p < .001$; NLA: $F(3,210) = 1076$, $p < .001$; NEU: $F(3,210) = 3003$, $p < .001$; PLA: $F(3,210) = 3640$, $p < .001$; PHA: $F(3,210) = 3338$, $p < .001$). As shown in Figure S2A, within the NHA condition, both groups tended to rate faces as less negative than pictures and words, and in the NLA condition, both groups rated words as more negative than pictures and faces. In the neutral condition, both groups rated faces as more negative than pictures and words. In the PHA and PLA conditions, both groups rated pictures as less positive than faces and words.

A.





B. **Figure S2.** Behavioral valence and arousal ratings by stimulus type

For the arousal ratings (Figure S2B), there was a significant main effect of stimulus ($F(2, 140) = 23.41$, $p < .001$) and a significant stimulus-by-condition interaction ($F(8, 560) = 5.03$, $p < .001$). There were no significant group-by-stimulus ($F(2,140) = 0.36$, $p > .70$) or group-by-condition-by-stimulus ($F(8,560) = 0.76$, $p > .64$) interactions. Overall, words were rated as most arousing, faces were rated as least arousing, and pictures were intermediate. Simple effects tests revealed a significant effect of stimulus within each condition (NHA: $F(3,210) = 517.12$, $p < .001$; NLA: $F(3,210) = 797.53$, $p < .001$; NEU: $F(3,210) = 2427.5$, $p < .001$; PLA: $F(3,210) = 810.03$, $p < .001$; PHA: $F(3,210) = 591.5$, $p < .001$) and a significant effect of condition within each stimulus (faces: $F(5,350) = 1076$, $p < .001$; pictures: $F(5,350) = 872.55$, $p < .001$; words: $F(5,350) = 702.78$, $p < .001$). As shown in Figure S2B, in the negative conditions, both groups tended to rate faces as less arousing than pictures or words, and in the positive conditions, both groups tended to rate words as more arousing than pictures and faces. Arousal ratings to neutral stimuli did not differ between stimulus types in either group.

Effects of Stimulus Type: fMRI analysis: Table S3 lists regions that showed a significant interaction between stimulus type and each of the three contrasts at the whole-brain level, along with the results of follow-up analyses on the average activity within these regions. The valence and valence X arousal contrasts identified a similar set of regions, most of which showed within-stimulus effects only for pictures, though a few also showed significant effects for faces. The arousal contrast identified a number of regions driven primarily by pictures, words, or both. As mentioned in the main text, no region was identified that showed a significant stimulus X group interaction for any of the contrasts at the whole-brain

level. To further probe for an interaction with group, we performed simple effects tests looking for an effect of group within each stimulus type. With one exception, none of these regions showed a significant effect of group within any one stimulus type. Similarly, when we looked for effects of group within each stimulus type, we found that most of the regions showed a significant effect of stimulus type within each group individually.

Brain Region	BA	Talairach Coordinates	# Voxels	*Z*	Stim X Group Interaction	Within-stimulus effects			Within-stim group effects			Within-group stim effects	
						Pictures	Words	Faces	Pictures	Words	Faces	CON	SCZ
Valence Contrast:													
L Cerebellar Declive	-	-30, -57, -15	211	4.92	NS	****	NS	NS	NS	NS	NS	***	****
R Cerebellar Declive	-	33, -52, -11	346	5.16	NS	****	NS	NS	NS	NS	NS	***	****
L Middle Occipital Gyrus	19	-41, -80, 5	313	4.77	NS	****	NS	NS	NS	NS	NS	***	****
R Middle Temporal Gyrus	39	45, -73, 14	139	4.61	NS	****	NS	NS	NS	NS	NS	*	****
R Inferior Frontal Gyrus	9	43, 7, 30	50	3.93	NS	****	NS	****	NS	NS	NS	***	*
L Angular Gyrus	39	-50, -71, 33	37	4.07	NS	NS	*	****	NS	NS	NS	NS	****
R Cuneus	19	29, -85, 31	48	4.50	NS	****	NS	NS	NS	NS	NS	*	****
Arousal Contrast:													
L Middle Temporal Gyrus	39	-52, -64, 11	522	5.63	NS	****	****	NS	NS	NS	NS	****	****
L Inferior Frontal Gyrus	47	-42, 16, -4	116	4.88	NS	***	****	NS	NS	NS	NS	***	****
R Inferior Frontal Gyrus	47	36, 16, -6	38	4.58	NS	***	***	NS	NS	NS	NS	**	***
L Inferior Frontal Gyrus	9	-49, 16, 24	128	4.83	NS	NS	****	NS	NS	NS	NS	*	****
R Middle Occipital Gyrus	19	51, -75, 7	46	4.34	NS	****	NS	*	NS	NS	NS	***	*
L Precuneus	7	-3, -60, 32	295	4.64	NS	****	NS	NS	NS	NS	NS	***	****
L Superior Frontal Gyrus	9	-4, 51, 26	32	4.22	NS	****	NS	NS	NS	NS	NS	NS	****
L Superior Frontal Gyrus	8	-4, 16, 49	358	4.74	NS	**	***	NS	NS	NS	NS	**	****
L Paracentral Lobule	31	-2, -21, 45	47	4.21	NS	****	NS	NS	NS	NS	NS	NS	****
L Middle Frontal Gyrus	6	-32, 3, 59	55	4.11	*	*	***	NS	*	NS	NS	NS	****
Valence X Arousal Contrast:													
L Cerebellar Declive	-	-30, -56, -14	108	4.76	NS	****	NS	NS	NS	NS	NS	***	***
R Fusiform Gyrus	37	36, -55, -11	168	4.93	NS	****	NS	NS	NS	NS	NS	***	****
L Inferior Temporal Gyrus	19	-45, -76, 0	158	4.51	NS	****	NS	NS	NS	NS	NS	****	***
R Middle Temporal Gyrus	39	46, -71, 16	133	4.63	NS	****	NS	NS	NS	NS	NS	*	****
R Inferior Frontal Gyrus	9	43, 8, 31	45	3.89	NS	****	NS	****	NS	NS	NS	***	*
R Cuneus	19	29, -85, 32	42	4.46	NS	****	NS	NS	NS	NS	NS	*	****

*p<.05; **p<.01; ***p<.005; ****p<.001

Table S3. Regions showing a significant effect of stimulus for each contrast

L, left; R, right; BA, Brodmann area; NS, not significant; CON, controls; SCZ, schizophrenia

Whole-brain Correlation Analyses: Table S4 shows the results of the whole-brain correlation analyses between anhedonia scores and activation in the valence, arousal, and valence X arousal contrasts. No regions survived the threshold in controls, but several regions were significant in patients.

Table S4. Whole-brain correlation results in individuals with schizophrenia

Contrast	Anhedonia Measure	Talairach Coordinates	Region name	Brodmann Area	# Voxels	*r*	*z*
Arousal	Chapman social anhedonia	-3, 63, 0	L Superior Frontal Gyrus	10	89	0.68	4.81
Valence	Chapman physical anhedonia	-49, 17, -19	L Superior Temporal Gyrus	38	30	-0.56	-3.78
Valence-by-arousal	Chapman physical anhedonia	38, -76, -23	R Cerebellar Tuber	…	50	-0.56	-3.78
		13, -10, -9	R Amygdala	…	45	-0.63	-4.37
Arousal	Chapman physical anhedonia	8, -18, -13	R Midbrain	…	88	0.58	3.90
Valence-by-arousal	SANS global anhedonia	-28, -59, -21	L Cerebellar Culmen	…	59	-0.64	-4.42
		0, 8, -14	Subcallosal Gyrus	25	50	0.52	3.45
		59, -53, -13	R Inferior Temporal Gyrus	20	38	0.59	4.02
Arousal	SANS global anhedonia	-2, 63, 0	L Superior Frontal Gyrus	10	87	0.61	4.17

L, left; R, right; SANS, Scale for the Assessment of Negative Symptoms

Medication Effects

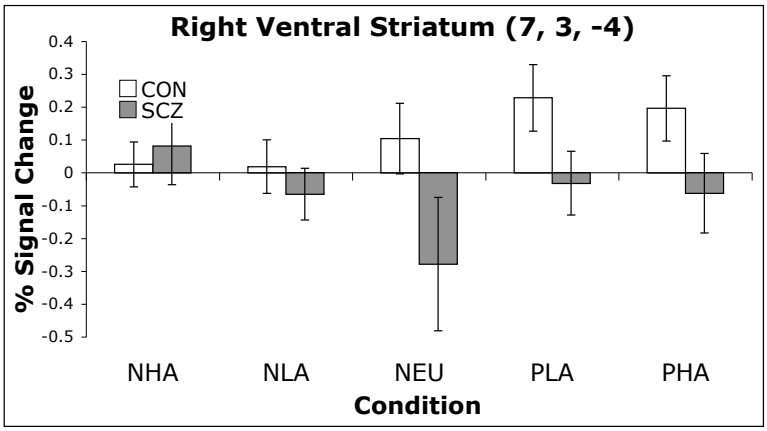


Figure S3. Functional activation in right ventral striatum in controls and in patients taking atypical antipsychotics only (excluding risperidone)

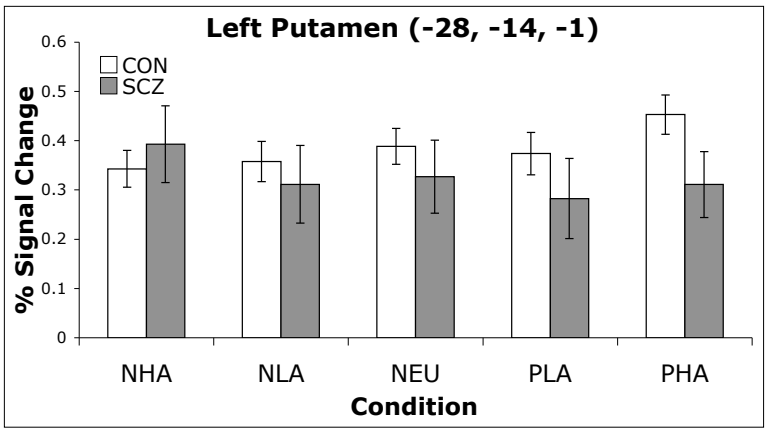


Figure S4. Functional activation in left putamen in controls and in patients taking atypical antipsychotics only (excluding risperidone)

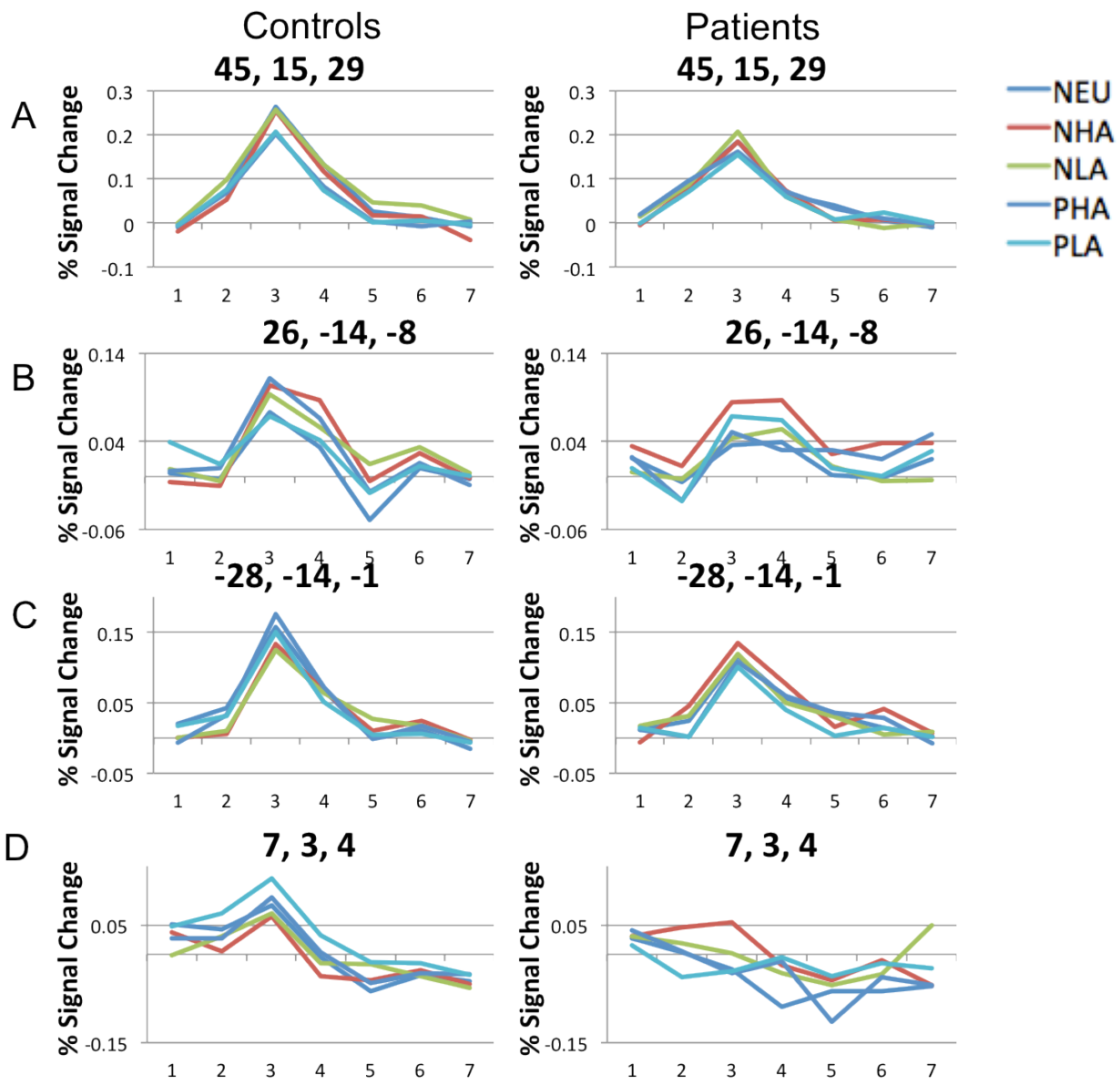
Schizophrenia versus Schizoaffective Disorder: Our sample included individuals with both schizophrenia and schizoaffective disorder. There has been debate in the literature as to whether or not these represent similar or different disorders. We chose to include individuals with schizoaffective disorder on the basis of a large body of literature suggesting that it is inappropriate to treat schizoaffective disorder as a distinct structural diagnosis. Critical reviews and meta-analyses of the literature in neuropsychology,

neuroimaging, molecular neurobiology, and genetic epidemiology have consistently failed to find categorical differences between schizoaffective disorder, schizophrenia, and bipolar disorder, leading several authors to conclude that the current diagnostic structure is somewhat artificial and a dimensional or spectrum approach to psychotic and affective disorders would be more appropriate (Abrams et al 2008; Cheniaux et al 2008; Lake and Hurwitz 2006; Malhi et al 2008; Peralta and Cuesta 2008). For this reason, we felt it was not necessary to exclude participants with a diagnosis of schizoaffective disorder from our study.

However, to be sure that diagnosis did not have an effect on the data in our sample, we conducted supplemental analyses in which we excluded patients with a diagnosis of schizoaffective disorder from our final sample. The major findings of the study remained unchanged. Behaviorally, the group X condition interactions remained significant for both the valence and arousal ratings, and post-hoc analyses revealed the same group differences within individual conditions as in the full sample. In the fMRI analyses, the group differences in the valence contrast in right ventral striatum and in the valence X arousal contrast in left putamen remained significant, and the negative correlations between physical anhedonia and bilateral amygdala activity remained significant. The negative correlation between right ventral striatal activity in the valence X arousal contrast and physical anhedonia dropped to trend level (from r = -.36, $p < .04$ to r = -.320, $p < .08$), possibly due to the reduction in power. We feel that these results show that the diagnosis of schizophrenia vs. schizoaffective disorder did not influence the outcome of our study, and therefore justify inclusion of these patients in our sample.

Example Timecourses

Figure S5: Example Timecourses. (A) Timecourses for a region in DLPFC showing a significant valence effect. (B) Timecourses for a region in right amygdala showing an arousal effect. (C) Timecourses for a region in left putamen showing group differences in the valence effect. (D) Timecourses for a region in right ventral striatum showing group differences in the valence effect.


Controls
Patients
A
45, 15, 29
B
26, -14, -8
C
-28, -14, -1
D
7, 3, 4
% Signal Change
NEU
NHA
NLA
PHA
PLA

Chapter 3.

Pavlovian reward prediction and receipt in schizophrenia: Relationship to anhedonia

This chapter was published in the May, 2012 issue of PLOS One. My contributions to this paper included data collection, analysis, and writing.

Reference: Dowd EC & Barch DM. Pavlovian reward prediction and receipt in schizophrenia: relationship to anhedonia. PLoS ONE (2012) vol. 7 (5) pp. e35622.

Abstract

Reward processing abnormalities have been implicated in the pathophysiology of negative symptoms such as anhedonia and avolition in schizophrenia. However, studies examining neural responses to reward anticipation and receipt have largely relied on instrumental tasks, which may confound reward processing abnormalities with deficits in response selection and execution. 25 chronic, medicated outpatients with schizophrenia and 20 healthy controls underwent functional magnetic resonance imaging using a Pavlovian reward prediction paradigm with no response requirements. Subjects passively viewed cues that predicted subsequent receipt of monetary reward or non-reward, and blood-oxygen-level-dependent signal was measured at the time of cue presentation and receipt.

At the group level, neural responses to both reward anticipation and receipt were largely similar between groups. At the time of cue presentation, striatal anticipatory responses did not differ between patients and controls. Right anterior insula demonstrated greater activation for nonreward than reward cues in controls, and for reward than nonreward cues in patients. At the time of receipt, robust responses to receipt of reward vs. nonreward were seen in striatum, midbrain, and frontal cortex in both groups. Further, both groups demonstrated responses to unexpected versus expected outcomes in cortical areas including bilateral dorsolateral prefrontal cortex. Individual difference analyses in patients revealed an association between physical anhedonia and activity in ventral striatum and ventromedial prefrontal cortex during anticipation of reward, in which greater anhedonia severity was associated with reduced activation to money versus no-money cues. In ventromedial prefrontal cortex, this relationship held among both controls and patients, suggesting a relationship between anticipatory activity and anhedonia irrespective of diagnosis. These findings suggest that in the absence of response requirements, brain responses to reward receipt are largely intact in medicated individuals with chronic schizophrenia, while reward anticipation responses in left ventral striatum are reduced in those patients with greater anhedonia severity.

Introduction

The role of reward processing in the pathophysiology of schizophrenia has garnered significant attention in recent years. Aberrant reward processing has been implicated in both positive (Heinz and Schlagenhauf 2010; Kapur 2003) and negative (Barch and Dowd 2010; Gold et al 2008) symptomatology, and advances in neuroimaging techniques have allowed new insights into the mechanisms of reward processing that may be disrupted in this illness. One such process is reward prediction, or the ability to anticipate a reward when presented with a predictive cue. Reward prediction is strongly associated with dopaminergic activity in the midbrain and striatum (O'Doherty 2004; Schultz 2001), which is thought to be dysregulated in schizophrenia (Guillin et al 2007). If disruptions in reward signaling prevent predictive stimuli from taking on the appropriate significance, they could contribute to important negative symptoms of schizophrenia such as decreased motivation and anhedonia (a reduced ability to experience pleasure) (Barch and Dowd 2010; Gold et al 2008; Ziauddeen and Murray 2010). Here, we examine this possibility using a Pavlovian reward prediction task to examine functional activity during reward anticipation and receipt in schizophrenia and its relationship to symptoms of anhedonia and amotivation.

A number of previous neuroimaging studies have examined reward prediction in schizophrenia. Studies using monetary incentive delay paradigms have shown reduced ventral striatal responses to reward-predictive cues in patients who are unmedicated (Juckel et al 2006b), or taking typical, but not atypical, antipsychotics (Juckel et al 2006a). Notably, this reduction in anticipatory activation was associated with negative symptom severity. Several studies in patients taking atypical antipsychotics have shown intact striatal anticipatory activation (Simon et al 2010; Walter et al 2009; Waltz et al 2010), though some of these studies also demonstrated negative correlations between ventral striatal activation and negative symptoms (Simon et al 2010; Waltz et al 2010).

Work examining brain responses to reward receipt has also revealed alterations in schizophrenia. Some studies have shown blunted striatal reward responses or prediction errors (responses to outcomes that do not match expectation) (Gradin et al 2011; Koch et al 2010; Murray et al 2008; Schlagenhauf et al 2009; Waltz et al 2009), while others have shown intact responses (Simon et al 2010; Waltz et al 2010). Notably, several of these studies also found inverse relationships between striatal responses to reward receipt and negative or depressive symptoms (Simon et al 2010; Waltz et al 2009; Waltz et al 2010). In addition, abnormal outcome or prediction error responses have been reported in cortical regions including

insula (Gradin et al 2011) and medial (Murray et al 2008; Schlagenhauf et al 2009; Walter et al 2009; Waltz et al 2010), ventrolateral (Walter et al 2010; Walter et al 2009), and dorsolateral (Corlett et al 2007; Waltz et al 2010) prefrontal cortex. In several of these studies, attenuated cortical prediction errors or outcome responses were associated with increased severity of either positive (Corlett et al 2007; Gradin et al 2011; Murray et al 2008) or negative (Walter et al 2010; Waltz et al 2010) symptoms.

Importantly, the literature examining reward processing in schizophrenia has largely relied on instrumental learning tasks, in which rewards must be earned via correct and/or rapid response execution. In these tasks, the ability to anticipate a reward depends upon the ability to earn the reward by responding appropriately. This requires not only reward prediction, but also action selection and response execution, any of which may be impaired in schizophrenia. Here, we examined reward prediction in schizophrenia in the absence of requirements for response selection and execution. Using a passive Pavlovian paradigm, we examined functional activation in response to rewarding stimuli and to predictive cues that had been associated with rewards based on pre-scan instructions.

Previous work using aversive Pavlovian conditioning has revealed abnormal brain responses among individuals with schizophrenia. Using a task in which colored cues were associated with affectively negative or neutral pictures, Romaniuk et al (Romaniuk et al 2010) demonstrated reduced responses to aversive cues in bilateral amygdala, as well as decreased prediction error responses in the midbrain. Further, inappropriate midbrain activation to neutral stimuli correlated with delusional symptom severity. Similarly, work by Jensen et al (Jensen et al 2008) using aversive noise stimuli revealed increased right ventral striatum activation to neutral cues in patients. These studies suggest that even in the absence of response requirements, brain responses to neutral cues in an aversive context may be augmented among individuals with schizophrenia. In addition, two studies have examined functional activity using Pavlovian paradigms with appetitive rewards in schizophrenia (Morris et al 2012; Waltz et al 2009). Waltz et al used a timing-sensitive paradigm to examine anticipation and receipt of primary reward (juice), and found reduced positive (but not negative) prediction errors and reward responses in schizophrenia in widespread regions throughout the brain. However, integrating these results with others in the literature is challenging because the timing-sensitive paradigm and primary reward differ greatly from the tasks typically used in instrumental studies. Morris et al used a Pavlovian prediction error task to

examine responses to expected and unexpected rewards and omissions. In this study, ventral striatal responses in patients were intact for reward receipt vs. omission, but failed to differentiate between expected and unexpected rewards. However, this study did not examine brain activity at the time of cue presentation. Here, we used a Pavlovian monetary reward prediction task to examine whether functional activation during reward anticipation and receipt is altered in schizophrenia even in the absence of response requirements, and whether it relates to symptoms of anhedonia and amotivation.

Methods and Materials

Participants: Participants were 29 stable outpatients with DSM-IV schizophrenia or schizoaffective disorder and 22 healthy controls with no personal or family history of psychosis. All patients were taking antipsychotic medications, which were stable for at least two weeks. Participants were group matched on sex, age, parental education, handedness (Oldfield 1971), and smoking status. Inclusion criteria were 1) age 18-50 years and 2) ability to give informed consent. Exclusion criteria were 1) DSM-IV substance abuse or dependence within the past 6 months (except nicotine); 2) DSM-IV major depressive disorder or dysthymia in the past year; 3) past head injury with neurological sequelae and/or loss of consciousness; 4) DSM-IV mental retardation, and 5) any contraindication to MRI including pregnancy, claustrophobia, any metallic object in the body, etc. Participant diagnoses were based on a Structured Clinical Interview for DSM-IV-TR (First et al 2001) conducted by a Masters-level clinician. Clinical symptoms were rated using the Scales for the Assessment of Positive Symptoms (SAPS) (Andreasen 1983b) and Negative Symptoms (SANS) (Andreasen 1983a), and summarized using the following symptom domain scores (Andreasen et al 1995): 1) positive symptoms – hallucinations and delusions; 2) negative symptoms – alogia, anhedonia, avolition, affective flattening and attentional impairment; and 3) disorganization – bizarre behavior, positive thought disorder, and inappropriate affect. Anhedonia was assessed using the Chapman revised physical and social anhedonia scales (Chapman and Chapman 1978; Chapman et al 1976; Eckblad et al 1982). This study was conducted in accord with APA standards for ethical treatment of human subjects. Written informed consent was obtained from all participants, and all procedures were approved by the Washington University Human Research Protection Office. Participant demographic and clinical characteristics are shown in Table 1.

Materials and Tasks: All participants underwent fMRI while performing a Pavlovian reward prediction task. Subjects were presented with one of two visual cues (pink cross or green circle), one predicting receipt of 75¢ for the trial, and one predicting receipt of 0¢ for the trial. The cues were followed by their predicted outcome 75% of the time, and with the opposite outcome 25% of the time. Participants were informed of the cue-outcome associations before the scan, and were told that the cue usually, but not always, predicted its associated outcome. They were also told that they could keep any money they were awarded during the task. In each trial, the visual cue was presented for 10 seconds, followed by a symbol indicating the outcome (+75¢ or +0¢) for 4 seconds. Inter-trial intervals varied pseudorandomly between 4 and 8 seconds. Participants completed four runs of 16 trials each, for a total of 64 trials (32 per cue type). Participants were paid $25/hour for their time, and were awarded an additional $20 in reward money upon session completion.

Image Acquisition and Processing: Imaging data was acquired on a 3T Siemens TIM TRIO system with a 12-channel head coil. High-resolution T1 images (TE=3.16ms, TR=2400ms, 176 slices, 1X1X1mm voxels) and T2 images (TE=96ms, TR=5s, 48 slices, 1.02x1x3mm voxels) were acquired to aid in registration to a common atlas space. Functional images were collected in four runs of 182 frames each using an asymmetric spin-echo echo-planar sequence (TR=2000ms, TE=27ms, FOV=256mm, flip=90°, 33 slices). Functional runs acquired axial images parallel to the anterior-posterior commissure plane with $4mm^3$ isotopic voxels. Stimuli were presented using PsyScope on a G3 Macintosh, with each trial onset triggered directly by a pulse from the scanner. The MR data was normalized across runs by scaling whole-brain signal intensity to a fixed value and removing the linear slope on a voxel-by-voxel basis to counteract effects of drift (Bandettini et al 1993). The MR data was then aligned to correct for head motion using rigid-body rotation and translation correction algorithms (Friston et al 1994; Snyder 1996; Woods et al 1992), which provide estimated movement parameters used to evaluate movement differences between groups. We also compared signal-to-noise ratios (SNR=mean/variance) between groups (Barch et al 2002), and removed runs or participants with movement or SNR values not meeting predetermined criteria (Table S1). Of the 29 patients and 22 controls who underwent the experimental protocol, four patients and two controls were excluded for excessive head motion, yielding the final sample of 20 controls and 25 patients. The images were then resampled into $3mm^3$ voxels, registered to

Talairach space using 12-parameter affine transformations, and smoothed with a 6mm FWHM Gaussian filter.

fMRI Data Analysis: All functional data was analyzed using in-house software. Data analysis was conducted using general linear models (GLMs) (Friston et al 1995; Friston et al 1994; Worsley 1994), which included task-related regressors as well as nuisance regressors for linear trends within runs and baseline shifts between runs. Canonical hemodynamic response shapes (Boynton functions) were used to estimate cue- and receipt-related activation. Regressors included two cue types (money and no-money) and four outcome types (expected money, expected no-money, unexpected money, unexpected no-money). The parameter estimates from the GLMs for each subject were entered into ANOVAs using subject as a random factor. To identify regions in which activation related to anticipation of reward (cue-related activity), we performed a repeated measures ANOVA with cue type (money, no-money) as a within-subjects factor and group (schizophrenia, control) as a between subjects factor. To identify regions in which activation related to reward receipt, we used cue type (cue money, cue no-money) and receipt type (receive money, receive no-money) as within-subjects factors and group (schizophrenia, control) as a between-subjects factor. Cue type was included as a factor in analyses of receipt in order to evaluate potential prediction error effects, which would be expected to modulate responses to receipt according to whether the outcome was expected or unexpected.

These ANOVAs were used in voxelwise whole-brain and ROI analyses. Whole-brain analyses were corrected for multiple comparisons using a p-value/cluster size threshold of $p<.003$ (two-tailed) and 13 voxels. This correction factor was determined by Monte Carlo simulations to provide a whole-brain false-positive rate of $p<.05$ (Forman et al 1995; McAvoy et al 2001); an approach equivalent to that employed by the Alphasim program in the AFNI software package. Second, voxelwise ROI analyses were conducted within an a priori mask consisting of a network of regions implicated in reward processing, an approach equivalent to the “small volume correction” procedure in the SPM software package. This mask, developed by Beck et al (Beck et al 2010), consisted of regions that were hand-drawn in Talairach space on the basis of anatomical landmarks and previously published coordinates (Ahsan et al 2007; Jensen et al 2007; Kable and Glimcher 2007; Knutson et al 2001; Knutson et al 2003; Kringelbach and Rolls 2004; Nitschke et al 2006; O'Doherty et al 2003a; O'Doherty et al 2004), including the dorsal and ventral

striatum, ventral tegmental area, substantia nigra, amygdala, orbitofrontal cortex (OFC), ventromedial prefrontal cortex (VMPFC), and insula (Figure S1). This analysis was corrected for multiple comparisons using a combined p-value/cluster size threshold ($p<.01$ (two-tailed) and 19 voxels) determined using Monte Carlo simulations to provide $\alpha<.05$ for the whole ROI mask.

We also conducted correlation analyses between BOLD contrasts and anhedonia scores. In these analyses, Cue Money – No-money and Receive Money – No-money contrasts were created for each subject and correlated with Chapman physical and social anhedonia scores. These correlations were conducted voxelwise, and were corrected using the same small-volume and whole-brain correction procedures described above. To explore whether these relationships were unique to the patient group, regions demonstrating significant correlations were also examined within the control group.

In addition to the analysis using canonical hemodynamic response shapes, a separate analysis was conducted that estimated response magnitudes for each MR frame separately (e.g. (Miezin et al 2000)), allowing examination of response timecourses across both cue and receipt. This analysis was analyzed using ANOVAs with time as a factor, including timepoints 1-7 (0-12 seconds) in the analysis for Cue, and 8-14 (14-26 seconds) in the analysis for Receipt. This analysis yielded comparable results, and example timecourses are shown in Figure S5.

Results

Participant characteristics:

Participant demographic and clinical characteristics are presented in Table 1. The patient and control groups did not differ significantly on age, sex, race, smoking status, or parental education. The control group demonstrated significantly higher personal education than the patient group. With respect to anhedonia severity, individuals with schizophrenia demonstrated significantly higher scores than controls on both the Chapman physical ($t(38)=-4.06$, $p<.001$) and social ($t(38)=-2.20$, $p<.001$) anhedonia scales.

Table 1. Clinical and demographic characteristics.

	CON	SCZ	
Age	33.20 (9.44)	31.44 (9.31)	
Education (years)	15.03 (2.34)	12.60 (2.40)*	
Highest Parental Education (years)	14.00 (1.95)	13.96 (3.08)	
Sex (% Male)	70	72	
Race (% Caucasian)	60	52	
Smoking status (% Smokers)	25	40	
Past Major Depressive Disorder (%)	5	12	
Past Substance Dependence (%)	10	24	
Chapman Social Anhedonia	2.00 (1.61)	3.50 (2.50)*	
Chapman Physical Anhedonia	2.44 (1.76)	6.00 (3.35)*	
Duration of Illness (years)	-	13.93 (8.36)	
Antipsychotic Medication		**% Taking**	**Average Dose (mg)**
Fluphenazine decanoate	-	4	25
Haloperidal		4	10
Haloperidal decanoate	-	12	53.33
Risperidone		24	3.58
Aripiprazole	-	24	26
Paliperidone	-	12	8
Clozapine	-	8	250
Olanzapine	-	4	20
Quetiapine	-	24	266.67
Ziprasidone	-	8	125
Other Medication			
Antidepressant	-	28	
Mood Stabilizer	-	16	
Anticholinergic	-	28	
SAPS/SANS Positive	-	2.10 (2.40)	
SAPS/SANS Negative	-	2.95 (3.64)	
SAPS/SANS Disorganization	-	0.72 (0.77)	

* $p < .05$; CON = control, SCZ = schizophrenia, SD = standard deviation, SAPS/SANS = Scale for the Assessment of Positive/Negative Symptoms (Andreasen 1983)

fMRI results: Reward Anticipation

Results of the voxelwise ROI and whole-brain analyses are reported in Table 2. No regions demonstrated a significant main effect of cue. There was a significant main effect of group in VMPFC in the ROI analysis and significant main effects of group in VMPFC, cerebellum and left posterior cingulate in the whole-brain analysis, all of which demonstrated greater activation overall in controls than in patients. In addition, a significant Cue X Group interaction was seen in a region in right anterior insula in the ROI analysis. This region showed greater activity for no-money than money cues in controls

(F(1,19)=8.74, p<.009), and greater activity for money than no-money cues in patients (F(1,24)=8.15, p<.009). There were no significant Cue X Group interactions in the whole-brain analysis.

Table 2. Cue-related activation.

Effect	Analysis	Brain Region	Brodmann Area	Talairach Coordinates	Voxels	Z	Activation Pattern
Group	ROI	VMPFC	32/10	+0, +44, +8	93	3.63	CON > SCZ
	WB	R Cerebellum	-	+14, -57, -46	18	4.06	CON > SCZ
		R VMPFC	32/10	+1, +47, +8	58	3.95	CON > SCZ
		L VMPFC	32/10	-12, +44, +0	20	3.75	CON > SCZ
		L Posterior Cingulate	31	-2, -38, +34	23	3.74	CON > SCZ
CueXGroup	ROI	R Anterior Insula	13	+32, +15, +0	25	3.79	CON: No Money > Money; SCZ: Money > No Money

R = Right, L = Left, CON = control, SCZ = schizophrenia, VMPFC = Ventromedial prefrontal cortex. Z values represent mean activation across the region.

Given past findings with similar paradigms, we had expected to see greater activity for money than no-money cues in striatal regions among controls. Thus, to further explore the nature of the striatal responses within each group, we extracted the mean activation across voxels within caudate, putamen, and nucleus accumbens ROIs ((Mamah et al 2007); Figure S2) for money and no-money cues. Cue (Money, No-Money) X Group (Control, Schizophrenia) ANOVAs revealed a significant main effect of cue within bilateral caudate (left: F(1,43)=4.16, p<.05; right: F(1,43)=5.12, p<.03), with greater activation to money than no-money cues (Figure 1). Post-hoc paired t-tests revealed that this effect was driven largely by the patient group, which demonstrated significant activation to money cues and deactivation to no-money cues (left: t(24)=2.33, p<.03; right: t(24)=2.23, p<.04). Controls showed activation to both money and no-money cues, with no significant differences between cue types. No other regions demonstrated a main effect of cue, and there were no significant main effects of group or Cue X Group interactions in the striatal ROIs. We also wished to examine the possibility that stronger anticipatory responses were evident only in the later runs, which could have resulted if the stimulus-outcome associations that were formed based on pre-scan instruction were strengthened by experience during the early trials of the scan. To this end, we repeated our cue-related analyses after excluding the first BOLD run, in essence treating this run as a practice. This analysis did not reveal any additional reward-related regions showing significant main effects of cue or cue X group interactions.

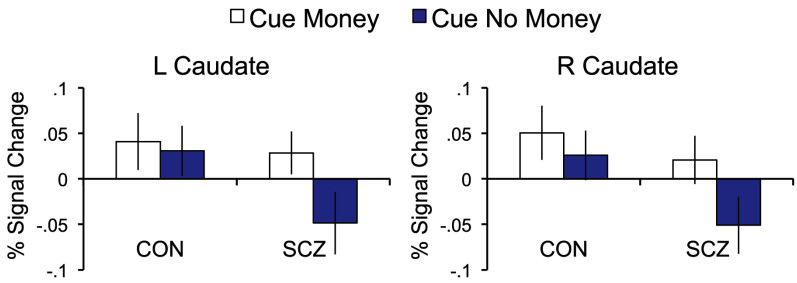


Figure 1. Cue-related activation in bilateral caudate ROIs. Activation shown is mean activation across voxels within regions of interest. Error bars represent standard error.

fMRI results: Reward Receipt

Results of voxelwise ANOVA analyses are presented in Table 3. The ROI analysis revealed a main effect of receipt in bilateral caudate, which showed greater activation for receipt of money than no-money. On whole brain analysis, a number of additional regions also showed this pattern, including bilateral DLPFC, dorsomedial PFC, left posterior parietal cortex, and several occipital and cerebellar regions (Figure 2). Post-hoc tests revealed that each group individually demonstrated greater activation for money than no-money receipt within each region at trend level or higher. In addition, several regions demonstrated a significant Cue X Receipt interaction (Figure 3). These included bilateral anterior insula and left ventrolateral prefrontal cortex, which activated more strongly for unexpected than expected outcomes (i.e. when money was cued but no-money received, or no-money cued and money received). Additional regions showing this pattern were also identified on whole-brain analysis. These included several regions commonly implicated in cognitive control and working memory (Owen et al 2005; Wager and Smith 2003), such as bilateral DLPFC, DMPFC, posterior parietal cortex, and anterior insula/frontal operculum. Post-hoc tests revealed that each of these regions demonstrated a significant cue X receipt interaction within each group separately (Table 3). To further explore the patterns of activity in these regions, we conducted paired t-tests comparing unexpected versus expected rewards and unexpected versus expected non-rewards within each group (table S2). All regions demonstrated significantly greater activation for unexpected than expected outcomes when collapsing across reward versus nonreward. When reward and non-reward were examined separately, all regions demonstrated greater activation for unexpected than expected rewards and for unexpected than expected non-rewards, though not every effect was significant in each group (details presented in Table S2). For example, the region in left

ventrolateral prefrontal cortex showed significantly greater activation for unexpected than expected rewards in controls, but not patients, and for unexpected than expected non-rewards in patients, but not controls (figure S3). These results suggest that the activation patterns driving the cue X receipt interactions were not identical between groups. However, no interactions with group were present, and both groups did show clear effects of unexpected versus expected outcomes. In addition to the regions showing main effects of receipt or cue X receipt interactions, a main effect of group was seen in right DLPFC, which demonstrated greater activity in controls than patients, as well as in left pre- and post-central gyri, which activated more strongly in patients than controls. Finally, one region in right inferior temporal gyrus demonstrated a Receipt X Group interaction wherein controls responded more strongly to no-money than money receipt, while patients responded more strongly to money than no-money receipt.

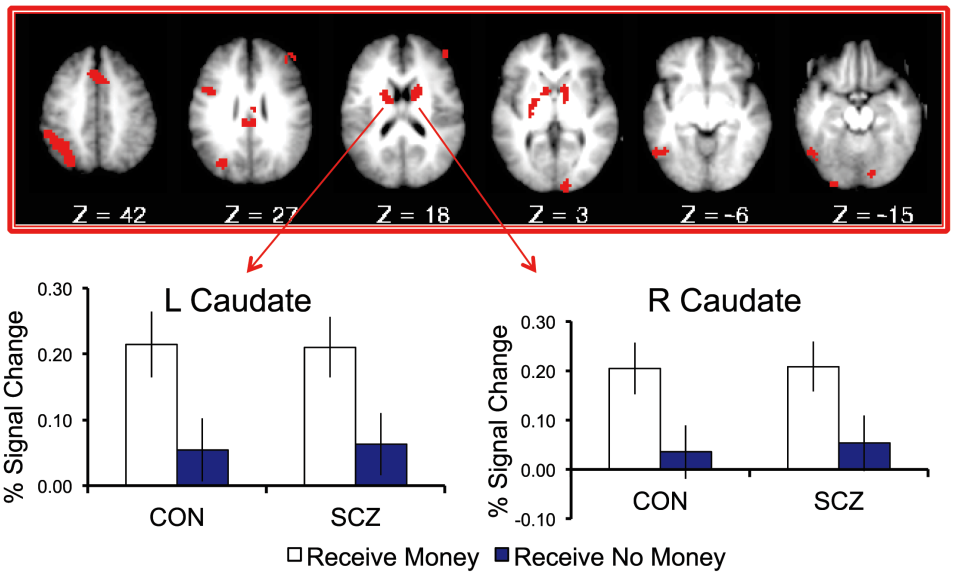


Figure 2. Regions demonstrating a significant main effect of receipt. All regions showed greater activation for receipt of money than no money. Both ROI results (threshold of p<.01 and 19 voxels within ROI mask) and whole-brain results (threshold of p<.003 and 13 voxels) are displayed. Graphs represent mean activation magnitudes across voxels within example regions among individuals with schizophrenia (SCZ) and controls (CON). Error bars represent standard error.

Table 3: Receipt-related activation

Effect	Analysis	Brain Region	Brodmann Area	Talairach Coordinates	# Voxels	Z	Activation Pattern	Post-hoc Tests: CON	SCZ
Receipt	ROI	L Caudate	-	-15, -4, +10	99	3.92	Money > No Money	p<.002	p<.02
		R Caudate	-	+12, +2, +13	74	3.94	Money > No Money	p<.001	p<.02
	WB	R Cerebellum	-	+1, -28, -41	40	4.09	Money > No Money	p<.02	p<.002
		L Cerebellum	-	-10, -46, -40	70	4.50	Money > No Money	p<.003	p<.002
		R Cerebellum	-	+11, -40, -43	15	3.70	Money > No Money	p<.03	p<.004
		L Cerebellum	-	-35, -62, -24	47	4.48	Money > No Money	p<.07	p<.001
		R Cerebellum	-	+53, -63, -24	15	4.38	Money > No Money	p<.005	p<.001
		R Cerebellum	-	+35, -65, -23	39	4.60	Money > No Money	p<.03	p<.001
		R Cerebellum	-	+12, -79, -19	22	4.14	Money > No Money	p<.03	p<.001
		L Middle Occipital Gyrus	37	-47, -62, -8	75	4.55	Money > No Money	p<.02	p<.001
		L Inferior Occipital Gyrus	18	-37, -84, -13	18	3.84	Money > No Money	p<.03	p<.002
		L Fusiform Gyrus	18	-25, -92, -12	20	4.12	Money > No Money	p<.02	p<.002
		R Cuneus	17	+11, -96, +1	40	3.86	Money > No Money	p<.06	p<.001
		L Thalamus	-	-12, -7, +10	107	4.39	Money > No Money	p<.001	p<.004
		R Caudate	-	+13, +1, +11	92	4.38	Money > No Money	p<.001	p<.005
		R Middle Frontal Gyrus	46	+43, +40, +22	44	4.18	Money > No Money	p<.002	p<.005
		L Inferior Frontal Gyrus	9	-42, +4, +26	55	3.61	Money > No Money	p<.03	p<.005
		Cingulate Gyrus	23	+0, -27, +28	65	4.14	Money > No Money	p<.006	p<.003
		L Angular Gyrus	39	-37, -57, +38	359	4.36	Money > No Money	p<.002	p<.003
		L Precuneus	7	-12, -74, +34	22	3.77	Money > No Money	p<.07	p<.002
		L Superior Frontal Gyrus	6	-2, +9, +48	141	4.07	Money > No Money	p<.02	p<.001
		L Precentral Gyrus	4	-44, -13, +42	16	3.59	Money > No Money	p<.1	p<.001
Group	WB	R Middle Frontal Gyrus	9	+43, +29, +33	14	3.76	CON > SCZ	-	-
		L Precentral Gyrus	4	-34, -28, +64	17	3.47	SCZ > CON	-	-
		L Postcentral Gyrus	3	-10, -32, +67	15	3.87	SCZ > CON	-	-
Cue X Receipt	ROI	L Ventrolateral PFC	47	-44, +22, -6	20	3.41	Unexpected > Expected	p<.001	p<.04
		R Anterior Insula	13	+34, +17, +4	57	4.12	Unexpected > Expected	p<.002	p<.005
		L Anterior Insula	13	-34, +16, +4	60	3.68	Unexpected > Expected	p<.005	p<.02
	WB	R Anterior Insula	13/45	+41, +21, +4	208	4.55	Unexpected > Expected	p<.002	p<.002
		L Anterior Insula	13/45	-40, +17, +0	89	4.21	Unexpected > Expected	p<.001	p<.009
		L Superior frontal Gyrus	9	-31, +47, +28	26	4.12	Unexpected > Expected	p<.006	p<.003
		L Posterior Cingulate	43	-1, -31, +26	25	3.39	Unexpected > Expected	p<.008	p<.04
		R Middle Frontal Gyrus	6	+42, +11, +42	137	4.65	Unexpected > Expected	p<.02	p<.001
		L Middle Frontal Gyrus	9	-44, +19, +32	18	3.84	Unexpected > Expected	p<.02	p<.004
		L Middle Frontal Gyrus	6	-39, +4, +34	14	3.65	Unexpected > Expected	p<.003	p<.02
		R Inferior Parietal Lobule	40	+44, -52, +40	151	4.40	Unexpected > Expected	p<.04	p<.001
		L Inferior Parietal Lobule	7	-39, -60, +42	139	4.11	Unexpected > Expected	p<.03	p<.001
		L Superior Frontal Gyrus	8	-1, +20, +48	67	4.17	Unexpected > Expected	p<.002	p<.007
Receipt X Group	WB	R Inferior Temporal Gyrus	37	+51, -70, +3	17	3.56	CON: No Money > Money SCZ: Money > No Money	-	-

R = Right, L = Left, ROI = Region of Interest analysis, WB = Whole Brain analysis, CON = control, SCZ = schizophrenia, PFC = prefrontal cortex. Z and p values represent mean activation across the region.

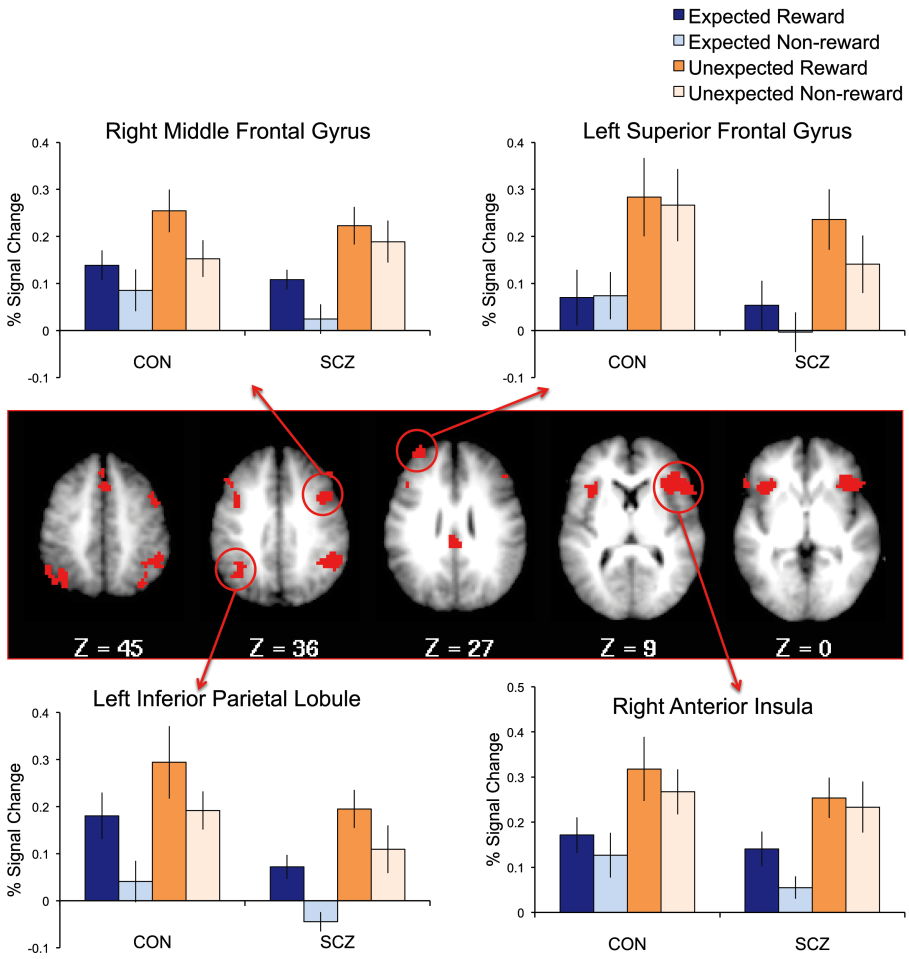


Figure 3. Regions demonstrating a Cue X Receipt interaction. All regions showed greater activity for unexpected than for expected outcomes. Both ROI results (threshold of p<.01 and 19 voxels within ROI mask) and whole-brain results (threshold of p<.003 and 13 voxels) are displayed. Graphs represent mean activation magnitudes across voxels within example regions among individuals with schizophrenia (SCZ) and controls (CON). Error bars represent standard error.

Correlations with Anhedonia

Cue-Related Activity Correlations. In the patient group significant negative correlations were found between Chapman physical anhedonia and the Cue Money – No-money contrast in regions in left ventral striatum and VMPFC (Figure 4a). As shown in Table 4, post-hoc analyses revealed that this relationship was primarily driven by activation in response to money cues, which correlated negatively with physical anhedonia, rather than by activation in response to no-money cues. This indicates that in patients, greater physical anhedonia is associated with less ventral striatal activity during the anticipation

of rewards. On whole-brain analysis, a similar pattern was observed in left inferior frontal gyrus. In VMPFC, the negative correlation was driven by activation to both the money cue, which correlated negatively with physical anhedonia, and the no-money cue, which correlated positively with physical anhedonia. Thus, individuals who were higher in anhedonia demonstrated both decreased responses to money cues and increased responses to no-money cues in VMPFC. To determine whether these relationships were also present among controls, we conducted correlation analyses in the control group between the Cue Money – No Money contrast and physical anhedonia scores within the ventral striatal and VMPFC regions that had reached significance in the patient group (Table 4). A significant negative correlation was found within the VMPFC region, indicating that greater physical anhedonia is associated with lower VMPFC anticipatory activity among both patients and controls. In the ventral striatal region, there was a negative relationship between anhedonia and cue-related activity in controls, but this relationship failed to reach significance. No significant regions were identified for social anhedonia.

Table 4: Correlations between Cue- and Outcome-Related Activity and Anhedonia Scores

Correlation		Analysis	Brain Region	Brodmann Area	Talairach Coordinates	Voxels	r	Money r	No-Money r	CON r
Cue Money - No Money	Physical Anhedonia	ROI	L Ventral Striatum	-	-17, 11, -1	27	-.720*	-.482*	0.408	-0.303
		ROI	L VMPFC	32	-3, 40, 4	20	-.652*	-.426*	.643*	-.493*
		WB	L VMPFC	24	-2, 33, 0	37	-.726*	-.502*	.670*	-0.459
		WB	L Inferior Frontal Gyrus	44	-51, 1, 20	14	-.726*	-.736*	0.148	0.1
Receive Money - No Money	Physical Anhedonia	WB	L Middle Frontal Gyrus	8	-30, 20, 38	23	.722*	0.351	-.457*	-0.186
		WB	L Superior Frontal Gyrus	6	-11, 24, 58	14	.762*	0.403	-.553*	-.627*
Receive Money - No Money	Social Anhedonia	WB	L Uncus	20	-15, -7, -34	31	-.772*	-.649*	.619*	-0.026
		WB	R Cerebellum	-	1, -47, -27	18	-.738*	-.569*	.477*	0.337

*p<.05. R = Right, L = Left, ROI = Region of Interest analysis, WB = Whole Brain analysis, CON = Control, SCZ = Schizophrenia. r values represent mean activation across the region.

Receipt Related Activity Correlations. At the time of receipt, significant positive correlations were seen between physical anhedonia and the Receive Money – No-money contrast in left superior and middle frontal gyri (Figure 4b). These correlations were driven primarily by responses to no-money receipt, which correlated negatively with physical anhedonia. In other words, patients who were higher in anhedonia showed less activation to no-money receipt in these regions. In addition, social anhedonia correlated negatively with activity in left uncus and right cerebellum in the Receive Money – No-money

contrast. These relationships were driven by activity in response to the receipt of both money and no-money; thus, patients with higher social anhedonia showed smaller responses in these regions to money receipt and larger responses to no-money receipt. In controls, only the left superior frontal gyrus showed a significant correlation, which was in the negative direction.

Figure 4

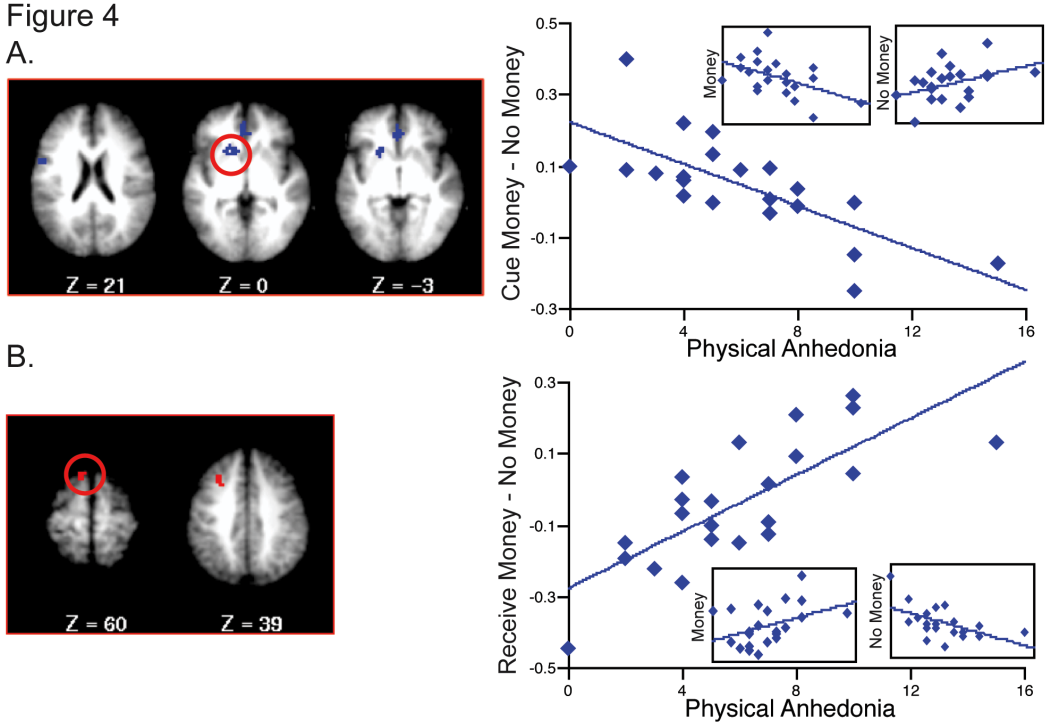


Figure 4. Regions demonstrating correlations between activation and anhedonia severity. (A) Results of voxelwise correlation between Chapman physical anhedonia score and the Cue Money – No Money contrast in patients, ROI analysis (threshold of p<.01 and 19 voxels within ROI mask). (B) Results of voxelwise correlation between Chapman physical anhedonia and the Receive Money – No Money contrast in patients, whole brain analysis (threshold of p<.003 and 13 voxels.) Graphs represent mean activation across the region.

Medication Analyses

To explore the possibility that antipsychotic medications may have influenced the results, we conducted correlations between medication dose in chlorpromazine equivalents (Gardner et al 2010)and cue- or receipt-related brain activation. First, we examined the regions that had shown correlations between cue- or receipt-related activity and anhedonia ratings, none of which showed significant correlations between antipsychotic medications and the relevant contrast (all *p* values > 0.16). Similarly, antipsychotic dosage failed to correlate significantly with physical or social anhedonia scores (all *p* values > 0.4). To look for more general medication effects, we also conducted voxelwise ROI and whole-brain

correlations between antipsychotic dose and activation for the money – no-money cue and receipt contrasts. There were no regions whose cue-related activity correlated significantly with antipsychotic dosage. For receipt, no regions within the ROI mask correlated significantly with medication dose. Whole-brain analysis revealed two regions that correlated positively with dose, such that higher doses were associated with greater responses to receipt of money versus no-money. One of these regions was within left prefrontal cortex (-30, 55, 0; 14 voxels) and the other fell within the lateral ventricles adjacent to caudate (3, 11, 12; 22 voxels).

In addition, given past results showing effects of typical vs. atypical antipsychotic medications on striatal anticipatory activation, we wished to determine to what extent our results were affected by medication type. To do this, we repeated our analyses with the 5 subjects taking typical antipsychotics removed (unfortunately, we had too few subjects taking typical antipsychotics to examine this group separately). All ANOVA and correlation results reported in Tables 2-4 remained significant at trend level or above, though some decreased slightly in significance perhaps reflecting the reduction in power. Further, no new correlations or group effects on striatal activation emerged at the time of cue presentation or receipt.

Discussion

The goal of this study was to examine cue- and receipt- related brain activation during Pavlovian reward prediction in schizophrenia, and to examine the relationship between this activation and symptoms of anhedonia and amotivation. Our results demonstrate few activation differences at the group level during reward anticipation and receipt. However, we observed that left ventral striatal and VMPFC activation during reward anticipation in schizophrenia was reduced in patients who are higher in anhedonia. These findings are consistent with a number of studies showing similar results in instrumental paradigms, consistent with the interpretation of these studies as reflecting abnormal reward prediction mechanisms in patients experiencing anhedonia and/or motivational deficits.

Cue-related activation: At the group level, we did not see striatal cue-related activity that differed significantly between groups. Upon ROI analysis, there was a main effect of cue in bilateral caudate with

greater activity for money than no-money cues, an effect that was driven by the patient group and was not significant in controls alone. We speculate that this relative lack of differential cue-related activity among controls as compared to published studies may reflect a difference between our passive Pavlovian paradigm and the instrumental paradigms typically used in reward prediction studies. Instrumental tasks require that information in the cue be used to plan and execute a correct motor response, which may enhance cue-related activation as compared to Pavlovian paradigms. Our task, modeled after O'Doherty et al 2002 (O'Doherty et al 2002), was designed to be completely passive in order to eliminate confounding variables associated with execution of a motor response. However, the O'Doherty study used primary reward (juice), and it is possible that the monetary rewards used here were less salient and therefore less capable of eliciting the expected cue-related activation. Similarly, it is possible that the monetary rewards were more salient for patients than controls given their lower socioeconomic status, which may underlie the finding of significant striatal anticipatory activation in patients but not controls.

Outside the striatum, we saw group differences in cue-related activity in the right anterior insula, such that patients showed greater activation to money than no-money cues, while controls showed the opposite pattern. The insula has been implicated in the detection of, and allocation of attention to, salient events (Menon and Uddin 2010), and has been shown to respond to unpleasant stimuli, including cues predicting negative (Jensen et al 2007) or low-expected-value (Rolls et al 2008b) outcomes. This is consistent with the pattern seen in controls, where insula activation was strongest to the less-pleasant, less-frequent cue. Patients, on the other hand, showed the opposite pattern, consistent with prior hypotheses (Kapur 2003) about an altered pattern of salience attribution in this group.

Receipt-related activation: Patients and controls showed similar activation at the time of reward receipt. A number of regions including bilateral caudate responded more strongly to money than no-money receipt in both groups, suggesting that brain responses to reward receipt were largely intact in this sample of patients. Similarly, we found greater responses to unexpected than expected outcomes in both groups in several regions of the cognitive control network (Cole and Schneider 2007), such as bilateral DLPFC, anterior insula, and posterior parietal cortex. These findings suggest that expectancy violations resulted in the recruitment of cognitive control regions in both patients and controls.

The results reported here are compatible with some, but not all, aspects of the results reported by Morris et al (Morris et al 2012). Both studies showed globally intact responses to reward receipt vs. omission, with both patients and controls showing reward responses in regions such as midbrain, insula, cingulate, and inferior frontal cortex. However, unlike Morris et al, we showed surprise responses in a number of cortical regions in both groups. In addition, while Morris et al reported group differences in ventral striatal prediction error activity, we failed to identify any regions whose activity pattern resembled a prediction error signal. This result is not surprising given the design of the task, which was intended to examine reward anticipation and receipt, not prediction error signaling. Because appreciable learning was not expected to occur and cue-outcome contingencies remained constant during the session, one would expect prediction error activity to be minimal.

Overall, our results demonstrate a pattern of largely intact responses to reward receipt throughout the brain at the group level among individuals with schizophrenia. The robust striatal responses to reward receipt among individuals with schizophrenia are consistent with recent studies demonstrating intact outcome responses in this region (Simon et al 2010; Waltz et al 2010). However, the pattern of intact cortical responses to reward receipt seen here contrasts with recent findings (Walter et al 2010; Walter et al 2009; Waltz et al 2010). One possible source of this discrepancy is that the referenced studies used instrumental tasks, which may engage cortical structures upon reward receipt to a greater extent than Pavlovian paradigms due to the requirements for updating future action plans. These results therefore suggest that responses to receipt of monetary reward per se may be intact in schizophrenia, and that reward processing deficits in this illness may therefore lie downstream in processes required to translate reward information into action plans. This conclusion is consistent with a large body of data suggesting intact hedonic responses to pleasant stimuli in schizophrenia, despite clinically evident deficits in motivation (Kring and Moran 2008). It is also consistent with the view that motivational deficits in schizophrenia are related to deficits in value representation (Gold et al 2008). In order to influence goal-directed behavior, information about reward receipt must be integrated and represented in a way that makes it available to guide value-based decision-making. Deficits in this process may not be evident during simple Pavlovian tasks, but may become evident in tasks where reward information must be used to guide future choices.

Relationships to Anhedonia: Correlation analyses revealed an inverse relationship between anticipatory activity in left ventral striatum and VMPFC and Chapman physical anhedonia scores in the patient group. This relationship is consistent with the findings of several reward prediction studies in schizophrenia showing that higher negative symptoms, particularly anhedonia and avolition, are associated with reduced striatal responses to reward-predicting cues (Juckel et al 2006b; Simon et al 2010). Along with Waltz et al (Waltz et al 2009), we have shown that this relationship is present in Pavlovian paradigms, suggesting that patients who are higher in anhedonia/avolition have larger deficits in reward prediction processes even in the absence of response requirements. Interestingly, the negative correlation seen in patients within the VMPFC region was also significant among controls, suggesting a relationship between activation to reward-predictive cues and individual differences in anhedonia irrespective of diagnosis. Previous studies in the literature examining non-clinical samples have identified relationships between anhedonia severity and brain activity in similar regions. Harvey et al (Harvey et al 2010) showed a negative correlation between anhedonia severity and activation in a rostral anterior cingulate cortex (rACC) region that overlaps with the VMPFC region identified here, and EEG studies by Wacker et al (Wacker et al 2009) showed increased resting delta activity (i.e. decreased resting activity) in rACC among more severely anhedonic individuals. These findings suggest that altered VMPFC activity is associated with anhedonia as a clinical dimension that is elevated in and associated with vulnerability to schizophrenia (Erlenmeyer-Kimling et al 1993; Kwapil 1998), but which crosses diagnostic boundaries.

We also observed relationships between anhedonia scores and activation at the time of receipt. In left frontal cortex, higher physical anhedonia was associated with decreased responses to nonrewarding outcomes. Because this study did not have a punishment condition and nonreward may have been considered a negative outcome in this context, this relationship may perhaps be thought of as a blunted response to negative outcomes among the more severely anhedonic patients. This finding is similar to past results, where anhedonia was found to relate to a reduced experience of both positive and of negative emotion (Dowd and Barch 2010).

When considering the relationships between anhedonia and brain activity in this study, it is worth noting that our sample had lower mean anhedonia scores than previous studies in the literature (Dowd and Barch 2010; Walter et al 2009). Given the relationship found between physical anhedonia and

anticipatory activity in patients, it is possible that a sample of patients who were higher in anhedonia may have shown differences at the group level between patients and controls, and that differences in anhedonia severity between samples may underlie some of the discrepant findings in the literature.

There is ample evidence from basic science research that hedonic experience and motivation are dissociable constructs subserved by different neurobiological systems (Berridge and Robinson 1998). However, the majority of clinical and self-report instruments available to assess anhedonia in schizophrenia, including the Chapman physical and social anhedonia scales used here, predate these findings and tend to contain items relevant to both constructs (Horan 2005). Here, we used the Chapman scales because of their wide range, which is well-suited to correlation analyses, as well as their widespread use and demonstrated reliability in the schizophrenia literature. It is important to note, however, that the non-specificity of these scales means that the relationships seen here may be driven by deficits in motivated behavior rather than by deficits in the experience of pleasure per se, consistent with the idea that a reduced ability to anticipate rewards may contribute to deficits in goal-directed behavior. Future studies using new clinical instruments that carefully distinguish between hedonic and motivational deficits, such as the Clinical Assessment Interview for Negative Symptoms scale (Horan et al 2011), will be required to dissociate the relationships of these two constructs to reward processing in schizophrenia.

Limitations: The passive task used in this study, while essential to address our experimental question about reward processing in the absence of response requirements, also confers an important limitation in that it did not allow the collection of behavioral data demonstrating attention to the task. This raises the possibility that our failure to find striatal reward anticipation responses among controls may have been driven by lack of attention to the stimuli. However, while we did not see significant *differential* responses between Money and No-money cues, we did see robust responses to cue presentation (irrespective of cue type) in regions including visual cortex, anterior insula, DLPFC, and posterior cingulate (figure S4). Further, we saw clear effects of money vs. no-money receipt in several brain regions within both groups, indicating that participants were attending to the outcome stimuli. Finally, we saw differential responses to unexpected versus expected outcomes, suggesting that participants were attending to the cues, developing cue-based expectancies, and reacting to expectancy violations upon outcome presentation. Together, we feel that these results provide reasonable evidence that participants

were indeed attending to the stimuli. We acknowledge that the degree of attentional engagement here was likely less than that required by active tasks, which is an additional reason to speculate that the relative lack of differential cue-related activity may reflect a difference between the Pavlovian paradigm used here and instrumental paradigms in the literature. An additional consequence of the passive task is that no behavioral measure of learning was available to determine whether stimulus-outcome associations were fully formed based on pre-scan instructions, or whether they were strengthened by experience during the early trials of the scan. Given the simplicity of the task, we expected instructed learning to be sufficient to establish cue-outcome associations and in-scanner learning to be minimal. However, to examine whether our results may have been influenced by any learning that took place during early trials, we examined anticipatory activity in runs 2, 3, and 4 only, after participants had had an additional 16 trials of training during run 1. This analysis did not reveal additional anticipatory activity or group differences in cue-related effects in striatum or other regions associated with reward processing. This suggests that although we are unable to assess the extent to which learning may have taken place in the scanner, it is not likely that additional pre-scan training would have resulted in stronger anticipation effects.

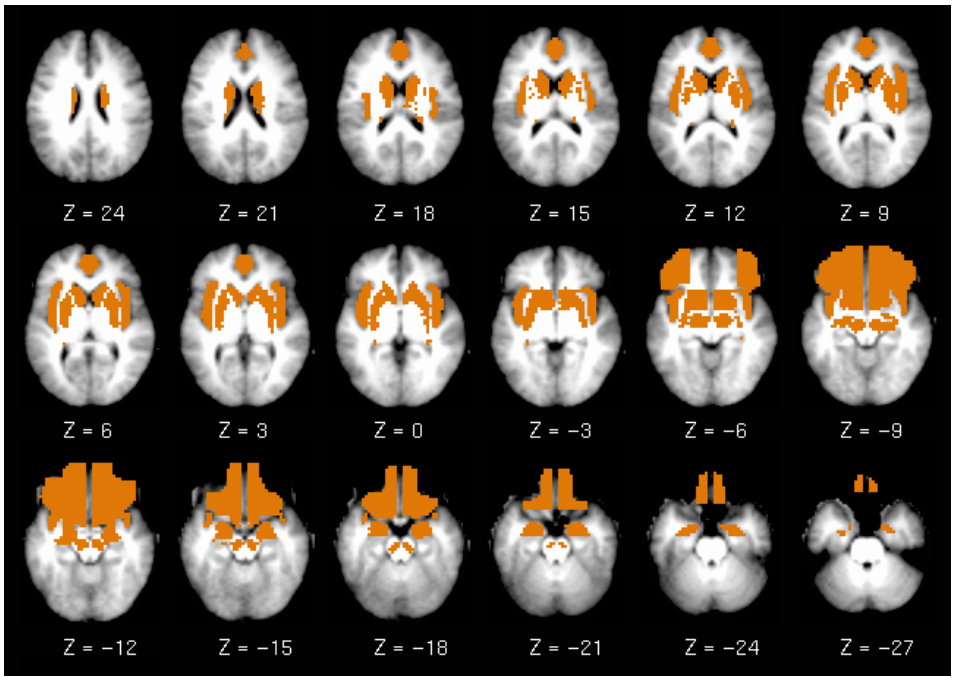
Another important limitation of this study is that patients were taking antipsychotic medications that block dopamine receptors, potentially affecting reward-related brain activation. However, our medication analyses failed to reveal significant relationships between antipsychotic dose and anhedonia score or BOLD signal within the regions of interest. These analyses suggest that the results reported here are not likely to be accounted for by medication effects, although studies including unmedicated patients are required to fully appreciate the extent to which medication may influence the processes examined here.

In sum, these findings suggest that while at the group level there were no differences between patients and controls in striatal activity during reward anticipation or receipt, those individuals with schizophrenia who were higher in anhedonia showed decreased striatal and VMPFC activity in anticipation of reward, even in the absence of requirements for response selection and execution. This suggests that the process of reward prediction may be abnormal in those patients experiencing the most severe anhedonia and motivational deficits.

Supplementary Materials

Regions of Interest

Figure S1: Regions of interest mask used in voxelwise ROI analyses:



Movement analysis

Four patients and two controls were excluded from analysis for excessive movement and/or low signal-to-noise ratio (SNR). Mean incremental (frame-to-frame) movement was computed for each run for each subject and used to compare movement between groups. SNR was computed by determining the ratio of the mean signal intensity to its standard deviation for each frame within each run, and mean SNR values were then calculated for each participant. For the final sample of 25 patients and 20 controls, mean SNR and incremental movement for each translation axis (x, y, z) and rotation axis (pitch, roll, yaw), as well as group t-tests results for each, are summarized in Table S1. The groups differed significantly only in movements along the y axis. Further, neither SNR nor movement along any axis correlated significantly with Chapman physical or social anhedonia score in patients. Together, these results indicate that the

groups were well matched for signal quality and that poor signal quality in patients is unlikely to contribute to the group or correlation results reported here.

Table S1. Incremental movement and signal-to-noise ratio

	Control		Schizophrenia		
	Mean	SE	Mean	SE	*p*
x	0.028	0.003	0.029	0.005	0.766
y	0.063	0.008	0.094	0.012	0.062
z	0.063	0.008	0.085	0.013	0.182
pitch	0.055	0.006	0.067	0.008	0.243
roll	0.024	0.002	0.029	0.003	0.212
yaw	0.021	0.002	0.026	0.004	0.333
SNR	276.34	72.83	277.90	76.11	0.945

SE, standard error; SNR, Signal-to-noise ratio
* represents p < .05

Cue-Related activity: Striatal ROI analyses

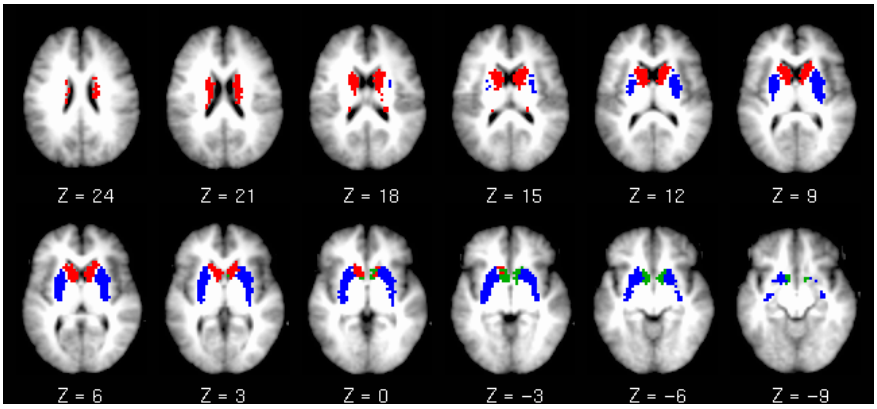


Figure S2: Caudate, Putamen, and Nucleus Accumbens ROIs used in mean-activation ROI analyses

Receipt-related activity: Post-hoc analyses for regions displaying CueXReceipt interactions

Brain Region	Brodmann Area	Talairach Coordinates	# Voxels	Paired t-tests					
				Unexpected vs Expected Outcomes		Unexpected vs Expected Reward		Unexpected vs Expected Non-reward	
				CON (p)	SCZ (p)	CON (p)	SCZ (p)	CON (p)	SCZ (p)
L VLPFC	47	-44, +22, -6	20	0.00072	0.032	0.0017	0.22	0.53	0.0073
R Anterior Insula	13	+34, +17, +4	57	0.0019	0.0049	0.1	0.4	0.0027	0.0004
L Anterior Insula	13	-34, +16, +4	60	0.0048	0.014	0.073	0.55	0.014	0.001
R Anterior Insula	13/45	+41, +21, +4	208	0.0015	0.0011	0.062	0.044	0.0084	0.00072
L Anterior Insula	13/45	-40, +17, +0	89	0.0007	0.0082	0.015	0.62	0.012	0.00023
L Superior frontal Gyrus	9	-31, +47, +28	26	0.0051	0.0028	0.038	0.0065	0.037	0.047
L Posterior Cingulate	43	-1, -31, +26	25	0.0073	0.035	0.08	0.13	0.01	0.05
R Middle Frontal Gyrus	6	+42, +11, +42	137	0.012	0.000097	0.078	0.0096	0.18	0.0016
L Middle Frontal Gyrus	9	-44, +19, +32	18	0.011	0.0036	0.06	0.083	0.057	0.03
L Middle Frontal Gyrus	6	-39, +4, +34	14	0.0025	0.015	0.1	0.037	0.07	0.12
R Inferior Parietal Lobule	40	+44, -52, +40	151	0.035	0.000091	0.46	0.066	0.026	0.000045
L Inferior Parietal Lobule	7	-39, -60, +42	139	0.024	0.00018	0.26	0.0051	0.028	0.0065
L Superior Frontal Gyrus	8	-1, +20, +48	67	0.0017	0.0064	0.03	0.72	0.026	0.0014

Table S2: Paired t-tests for unexpected versus expected rewards and non-rewards within patient and control groups

Figure S3: Receipt-related activation magnitudes for Left VLPFC

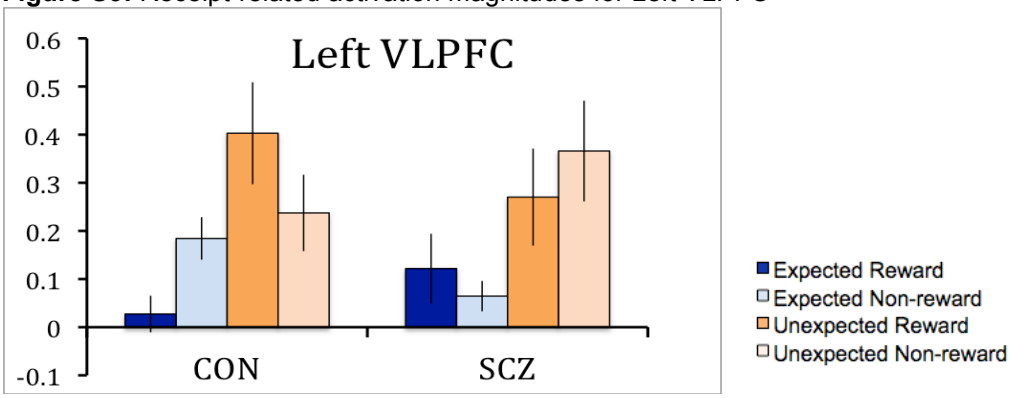
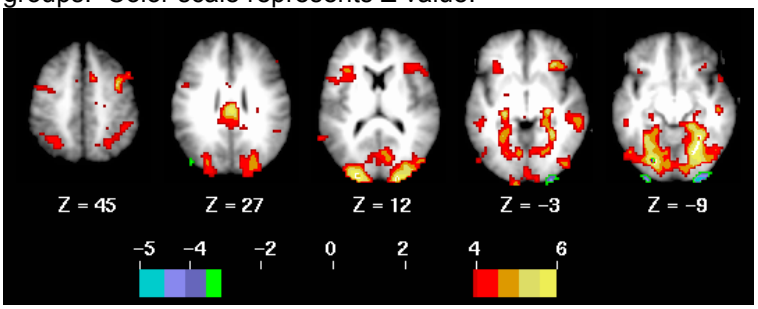
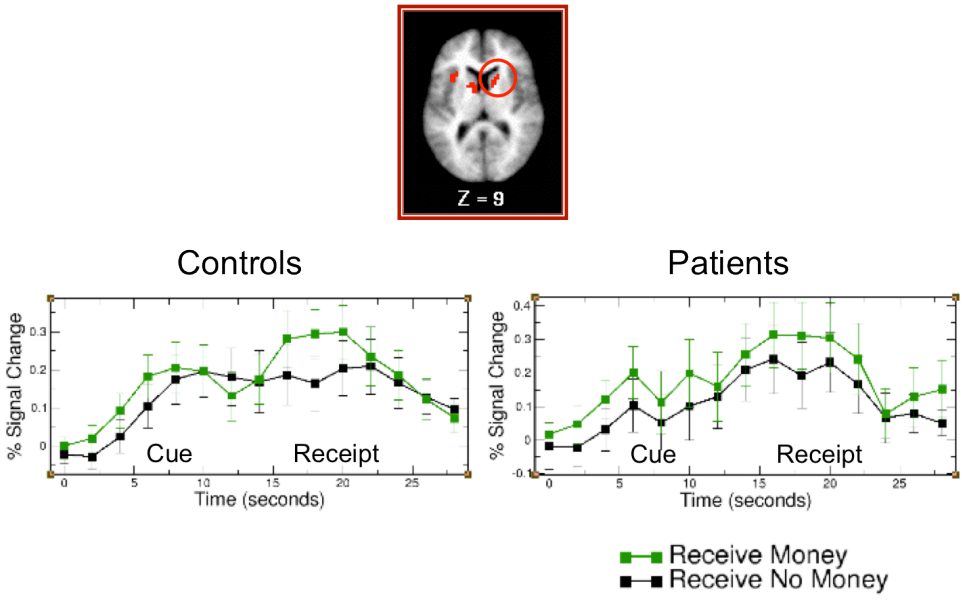


Figure S4: Cue-related activation for both cue types combined (money and no-money) across both groups. Color scale represents Z value.



Example Timecourses

Figure S5: Example timecourses across both Cue and Receipt for a right caudate region showing a main effect of receipt.



Chapter 4.

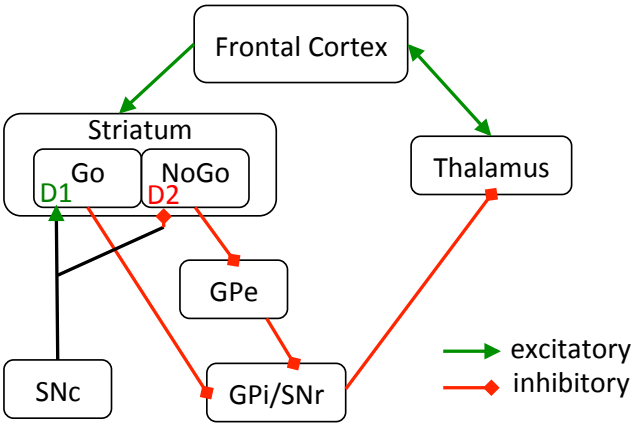
Probabilistic Reinforcement Learning in Schizophrenia

Introduction:

As discussed in Chapter 1, reinforcement learning (RL), the process by which actions that result in rewarding outcomes are strengthened and those that do not are suppressed, is an essential component of the translation of reward information into goal-directed behavior. Converging data from electrophysiology, neuroimaging, and computational modeling implicates the mesolimbic and nigrostriatal dopamine systems in RL, suggesting that these systems code reward prediction errors that gradually integrate outcomes over several trials (O'Doherty et al 2003b; Schultz et al 1997). This process is formally described by computational RL algorithms, which may be used to make predictions about data in neurophysiology and neuroimaging studies.

In addition to abstract RL algorithms, which describe information processing from a machine learning standpoint and do not make assumptions about physiology, several anatomically constrained neural network models have also been developed to describe the role of DA in RL (Cohen and Frank 2009). One such model (Frank et al 2004) emphasizes the separate contributions of D1Rs and D2Rs in the striatum to Go and NoGo learning. This model is founded upon the anatomy of the antagonistic "direct" and "indirect" pathways of the basal ganglia (Figure 1). Phasic dopaminergic activity in the substantia nigra pars compacta (SNc) can modulate activity in "Go" cells via D1Rs, which have excitatory effects, and in "NoGo" cells via D2Rs, which have inhibitory effects. When an action results in a positive prediction error, bursts of DA from SNc enhance direct pathway activity and inhibit indirect pathway activity, thereby driving long-term potentiation in the corticostriatal circuit and enhancing the representation of the chosen action in frontal cortex. Conversely, when a chosen action results in a negative prediction error, the resulting decrease in SNc dopaminergic activity inactivates the direct pathway and activates the indirect pathway, suppressing corticostriatal activity and weakening the cortical representation of the chosen response. In this way, associations between the stimuli and actions leading to positive outcomes are strengthened, while associations with actions leading to negative outcomes are weakened.

Figure 1: Frank model of “Go” and “NoGo” learning



The Frank model makes specific predictions about the effects of perturbations in dopaminergic function on RL, several of which have garnered empirical support. For example, the model predicts that Parkinson’s Disease patients taking L-DOPA will show enhanced positive RL and impaired negative RL as compared to unmedicated patients because greater dopamine availability should increase the efficacy of phasic firing, but decrease sensitivity to transient decreases in firing. Behavioral data from a probabilistic RL task allowing the separate assessment of learning from positive and negative outcomes supported both predictions (Frank et al 2004): while controls learned equally well from both positive and negative outcomes, there was a crossover effect in patients where medicated patients showed enhanced Go learning and impaired NoGo learning, while unmedicated patients showed impaired Go learning and enhanced NoGo learning. An analogous crossover effect was seen in a pharmacological study of healthy adults given a D2R agonist or antagonist (Frank and O'Reilly R 2006). Finally, a behavioral genetics study found that a polymorphism in the DARPP-32 protein, associated with striatal D1R function, was shown to modulate positive RL, while a polymorphism in the DRD2 gene, which affects postsynaptic D2R density in the striatum, modulated negative RL (Frank et al 2007). This data showing that the Frank model has made specific and supported predictions about RL in several cases of altered DA function makes it an ideal framework within which to examine RL in schizophrenia, which is associated with DA dysregulation (Goto and Grace 2007; Guillin et al 2007; Howes et al 2009).

Two studies using different tasks have examined Go and NoGo learning in medicated patients with schizophrenia (Waltz et al 2007), and found evidence of impaired Go learning but intact NoGo learning. The first study used a probabilistic stimulus selection task with transfer measures sensitive to Go vs

NoGo learning, on which patients demonstrated reduced performance on the Go learning measure, but not the NoGo learning measure. The second study obtained similar results, and also used the Frank neural network model of Go and NoGo learning to simulate behavior on the task. This simulation showed that modeling an increase in tonic dopamine coupled with a decrease in phasic dopamine, alterations consistent with current evidence about dopaminergic dysfunction in schizophrenia, reproduced the behavior seen in the patient group. These findings are consistent with the hypothesis that the effectiveness of phasic dopamine signals in response to positive feedback is reduced in schizophrenia, thereby impairing Go learning.

In addition to examining gradual basal ganglia-mediated learning, these studies also examined measures of win-stay/lose-shift activity during early learning, an indication of explicit learning thought to rely on PFC function. Both studies revealed decreases in this behavior, consistent with the hypothesis discussed in Chapter 1 that cortical explicit learning is more impaired in schizophrenia than gradual striatal-based learning. Importantly, in the studies discussed above, negative symptom severity correlated with decreases in early performance or win-stay/lose-shift behavior, but not with indicators of Go and NoGo learning impairment, suggesting that negative symptoms may be related to impairments in cortical rather than striatal learning. Further supporting this hypothesis are results from a recent study using two computational models, an actor-critic model thought to simulate striatal learning and a mixed actor-critic/Q-learning model thought to simulate learning from both striatal and cortical mechanisms. In this study, the behavior of patients who were high in negative symptoms was reproduced by the actor-critic model, while controls and patients with less severe negative symptoms were better described by the hybrid model. These results again suggest that impairments in cortical explicit learning may be related to negative symptoms in schizophrenia.

As discussed in Chapter 1, studies examining reinforcement learning in schizophrenia have revealed evidence for reduced striatal prediction error signaling, particularly for positive prediction errors (Gradin et al 2011; Morris et al 2012; Murray et al 2008; Waltz et al 2009), although normal striatal responses have also been reported (Walter et al 2009). However, these studies had a different emphasis than ours here, which was to probe striatal responses to determine whether they behaved consistently with prediction errors, not to understand the contribution of these prediction errors to learning. To this end, these studies

relied on simple associative learning tasks that minimized performance differences between patients and controls. In some cases, learning took place before the scan session began (Morris et al 2012; Waltz et al 2009), effectively probing prediction errors during reward-prediction rather than during reinforcement learning. While prediction errors, reward prediction, and reinforcement learning are all intimately related, we wished to determine whether any reductions in striatal prediction errors were associated with reinforcement learning impairments, especially in learning from positive outcomes, among individuals with schizophrenia. We also examined cortical activity during learning to determine whether impairments in striatal or cortical learning systems are impaired in schizophrenia, and whether these impairments are associated with anhedonia/avolition.

Methods and Materials

Participants: Participants were 49 stable outpatients with DSM-IV schizophrenia or schizoaffective disorder and 41 healthy controls with no personal or family history of psychosis. Both medicated and unmedicated patients were recruited from the community, and medication status and dose was required to have been stable for at least two weeks. Participants were group matched on sex, age, race, parental education, handedness (Oldfield 1971), and smoking status. Inclusion criteria were 1) age 18-50 years and 2) ability to give informed consent. Exclusion criteria were 1) DSM-IV substance abuse within the past year, or substance dependence within the past 2 years (except nicotine); 2) DSM-IV major depressive disorder or dysthymia in the past year; 3) any unstable or severe medical disorder; 4) past head injury with neurological sequelae and/or loss of consciousness; 5) DSM-IV mental retardation, and 6) any contraindication to MRI including pregnancy, claustrophobia, any metallic object in the body, history of heart rhythm abnormalities, and presence of a heart pacemaker. After undergoing the MRI protocol, 10 individuals with schizophrenia and 4 healthy controls were excluded for excessive movement (described below), and an additional patient was excluded for having more than 50% nonresponse trials, yielding a final sample size of 37 controls and 38 patients (35 medicated, 3 unmedicated). This study was conducted in accord with APA standards for ethical treatment of human subjects. Written informed consent was obtained from all participants, and all procedures were approved by the Washington

University Human Research Protection Office. Participants were compensated for their time at a rate of $25 per hour, plus an extra $20 in reward money.

Diagnosis and clinical assessment: Participant diagnoses were based on a Structured Clinical Interview for DSM-IV-TR (First et al 2001) conducted by a Masters-level clinician. Clinical symptoms were rated using the Scales for the Assessment of Positive Symptoms (SAPS) (Andreasen 1983b) and Negative Symptoms (SANS) (Andreasen 1983a), which were summarized using the following symptom domain scores (Andreasen et al 1995): 1) positive symptoms – hallucinations and delusions; 2) negative symptoms – alogia, anhedonia, avolition, affective flattening and attentional impairment; and 3) disorganization – bizarre behavior, positive thought disorder, and inappropriate affect. Clinician-rated symptoms of anhedonia and avolition were assessed with the SANS and the Brief Negative Symptom Scale (BNSS) (Kirkpatrick et al 2011). Given that anhedonia and avolition scores derived from the SANS correlate significantly with those derived from the BNSS, and that anhedonia and avolition scores load onto a single factor in both scales (Kirkpatrick et al 2011), we created a single score for clinician-rated symptoms of anhedonia and avolition by Z-scoring and summing across measures. Anhedonia and avolition were also assessed using the following self-report questionnaires, administered on a computer: The Revised Chapman Physical and Social Anhedonia Scales (Chapman and Chapman 1978; Chapman et al 1976), the Temporal Experience of Pleasure Scale (Gard et al 2007), the Snaith-Hamilton Pleasure Scale (Snaith et al 1995), and the Apathy Scale (Starkstein et al 1992), which were Z-scored and summed to create a composite measure of self-reported anhedonia/avolition. In addition the Wechsler Test of Adult Reading was administered as an estimate of premorbid IQ, and the Matrix Reasoning subtest of the Wechsler Adult Intelligence Test – III was administered as an estimate of current cognitive ability. Participants were required to pass a urine drug screen and breathalyzer test at the start of each session.

Materials and Tasks: The experimental paradigm was a modified version of the Probabilistic Stimulus Selection Task developed by Frank and colleagues (Figure 2) (Frank et al 2004). The task consisted of an acquisition phase, during which fMRI scanning took place, and a test phase that was completed outside the scanner. Stimuli consisted of grayscale drawings from the revised Snodgrass and Vanderwart object pictorial set (Rossion and Pourtois 2004), matched for luminance and contrast, visual complexity,

and object familiarity. Object-condition mappings were counterbalanced across subjects. During the acquisition phase, participants were presented on each trial with one of three pairs of stimuli ("AB", "CD", or "EF"), presented in pseudorandomized order, and were instructed to choose the stimulus in each pair that they find to be more often rewarded. Stimuli were displayed for 2000ms, during which the participant was required to choose one of the stimuli via button press. After a jittered inter-stimulus interval ranging from 2000-6000ms, feedback consisting of the words "Correct! +$" in green text, "Incorrect $0" in red text, or "Too Slow!" in red text were presented on screen for 2000ms. Subjects were told that for each "Correct" choice, they would win some extra money, with a total of up to $20 available to be won (in actuality, all subjects were paid an additional $20 upon completion). For stimulus pair AB, choice of A was rewarded 80% of the time, while B was rewarded 20% of the time; pair CD was 70:30, and pair EF was 60:40. Feedback was followed by an inter-trial interval jittered from 2000-6000ms. The stimuli were presented in 10 blocks of 36 trials (12 per stimulus pair), after which the test phase was administered outside the scanner. In this phase, the three original stimulus pairs were presented, along with 12 novel pairs in which the 3 original pairs were recombined. Each pair was presented 10 times, and participants were asked to choose the more frequently rewarded member of each pair based on the knowledge acquired earlier. No feedback was given at this stage. Trials were separated with 1000ms crosshairs triggered upon response, and no time limit was imposed on test-phase responses. As in previous studies using this task (Frank et al 2004), the recombined pairs from the test phase were used to calculate transfer measures indicative of learning from positive versus negative outcomes. Because A was the most highly reinforced stimulus during training, a "ChooseA" measure created by averaging performance on all novel pairs including A was used to indicate Go learning, and an analogous "AvoidB" measure to index NoGo learning.

Given pilot data indicating that individuals with schizophrenia had more difficulty understanding the task than controls, patients underwent a training session within the week prior to scanning where they completed 360 task trials with different stimuli. Both groups also completed a 12-trial practice session immediately before scanning.

Figure 2: Experimental Paradigm

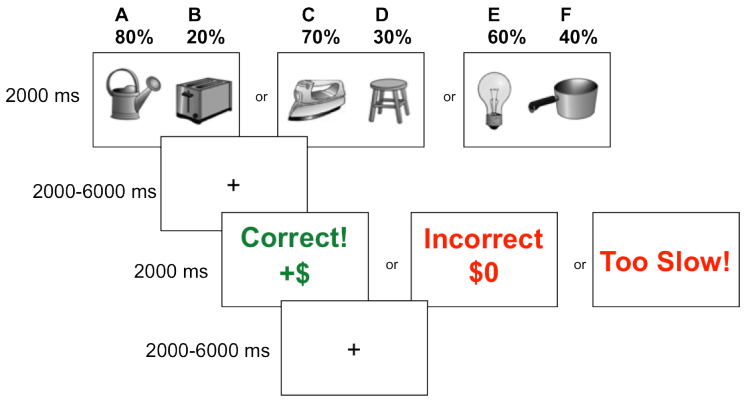


Image Acquisition and Processing: Imaging data was acquired on a 3T Siemens TIM TRIO system with a 12-channel head coil. High-resolution structural images were acquired using a sagittal magnetization-prepared rapid acquisition gradient echo (MP-RAGE) sequence (TR = 2.4s, TE = 3.08ms, inversion time = 1s, flip angle = 8°, 176 slices, 1 mm^3 voxels). Functional images were collected in 10 runs of 213 frames each using a gradient echo echo-planar sequence (TR=2030ms, TE=27ms, flip=90°, 36 slices). Functional runs acquired axial images parallel to the anterior-posterior commissure plane with $4mm^3$ isotopic voxels. Stimuli were presented using E-prime 2.0, with each trial onset triggered directly by a pulse from the scanner. The MR data was normalized across runs by scaling whole-brain signal intensity to a fixed value and removing the linear slope on a voxel-by-voxel basis to counteract effects of drift (Bandettini et al 1993). The MR data was then aligned to correct for head motion using rigid-body rotation and translation correction algorithms (Friston et al 1994; Snyder 1996; Woods et al 1992), which provide estimated absolute and frame-by-frame movement parameters used to evaluate movement differences between groups. We also compared mean voxelwise standard deviation values between groups (Table S1), and removed participants with movement or SD values not meeting predetermined criteria. The images were then resampled into $3mm^3$ voxels, registered to Talairach space using 12-parameter affine transformations, and smoothed with a 6mm FWHM Gaussian filter.

fMRI analysis: Statistical analysis of fMRI data was conducted based on the general linear model (GLM) (Friston et al 1995) using in-house software as described in (Miezin et al 2000). The general linear model for each subject included time as a 9-level regressor, made up of the 9 MR frames following each event. Response magnitude was therefore estimated at each time point separately, without assuming a specific shape for the hemodynamic response. Activation at the time of stimulus presentation and at the

time of feedback were modeled separately. At the time of stimulus presentation, 6 choice types were modeled (A, B, C, D, E, and F), and at the time of feedback, 12 feedback types were modeled (positive and negative feedback for each choice). Non-response trials were coded as a variable of no interest, and nuisance regressors for linear trends within runs and baseline shifts between runs were also included. Given that the choices for several subjects were not variable enough to code all event types, we also conducted a separate analysis in which all three stimulus pairs were combined. This analysis included only two choice types (high-probability (A/C/E), low-probability (B/D/F) and four feedback types (ACE/positive, ACE/negative, BDF/positive, BDF /negative). The parameter estimates from the GLMs for each subject, including time (the 9 time points of the response), were entered into ANOVAs using subject as a random factor. A significant main effect of time for a voxel or region indicates activation or deactivation, and a significant interaction of any other factor with time indicates that the hemodynamic response varies across that factor.

Activation at the time of stimulus presentation was evaluated using a repeated measures ANOVA with choice type (ACE/BDF) and time (1-9) within-subjects factors and group (schizophrenia, control) as a between subjects factor. Feedback-related activity was examined using a repeated measures ANOVA with choice type, feedback type (positive, negative), and time as within-subjects factors and group as a between subjects factor. Choice type was included as a factor in analyses of receipt in order to evaluate potential prediction error effects, which would be expected to modulate responses to receipt according to whether the outcome was expected or unexpected. These ANOVAs were used in voxelwise whole-brain and regionwise ROI analyses. Whole-brain analyses were corrected for multiple comparisons using a p-value/cluster size threshold of $p<.003$ (two-tailed) and 13 voxels, as determined by Monte Carlo simulations to provide a whole-brain false-positive rate of $p<.05$ (Forman et al 1995; McAvoy et al 2001). All regions demonstrating significant effects in whole-brain analyses were followed up with appropriate post-hoc tests conducted on the mean activation within the region. In addition, only regions demonstrating a significant main effect of time were reported. ROI analyses were conducted using mean activation within regions including bilateral caudate, putamen, and nucleus accumbens. These regions were defined anatomically (Mamah et al 2007) and were applied at the group level in Talairach atlas space. Significance levels in ROI analyses were False Discovery Rate corrected for multiple

comparisons using the Benjamini-Hochberg procedure (Benjamini and Hochberg 1995) to yield an alpha of .05 across all 6 regions.

We also conducted correlation analyses within the patient group between BOLD signal and anhedonia/avolition scores. In these analyses, activation at time points 4 and 5 (the peak of the hemodynamic response) was averaged and correlated with clinical and questionnaire-based anhedonia/avolition scores. Correlations were conducted for A/C/E choice events and for positive feedback events. These correlations were conducted using the same voxelwise whole-brain and regionwise ROI procedures described above. To explore whether these relationships were unique to the patient group, regions demonstrating significant correlations were also examined within the control group.

Model-based analyses. Behavioral data was modeled using a Q-learning algorithm with separate learning rates from positive feedback ("gains"; $\propto_G$) and negative feedback ("losses", $\propto_L$) (Frank et al 2007). This algorithm models subjects' choices by calculating a Q value, which is an estimate of expected reward value, for each stimulus (A-F). This value is modified on each trial according to the reward $r(t)$ received, where $r(t) = 1$ for positive feedback and $r(t) = 0$ for negative feedback. Expected value for stimulus i on trial $t + 1$ is calculated as follows:

$$Q_i(t+1) = Q_i(t) + \propto_G [r(t) - Q_i(t)]_+ + \propto_L [r(t) - Q_i(t)]_-$$

where $[r(t) - Q_i(t)]$, the reward received minus the reward expected, represents prediction error. In other words, the expected value for a given stimulus on a given trial is equal to its expected value on the previous trial plus an adjustment factor equal to prediction error times learning rate. Positive prediction errors (rewards that are higher than expected) are multiplied by the gain learning rate, and negative prediction errors (rewards that are lower than expected) are multiplied by the loss learning rate. Learning rates reflect the degree to which Q values are affected by reinforcement outcomes, with higher learning rates associated with larger changes in Q value on each trial. Action selection was modeled using a softmax logistic function, a standard stochastic decision rule that calculates the probability of choosing one stimulus over another given the expected reward values of both options:

$$P_A(t) = \frac{e^{\frac{Q_A(t)}{\beta}}}{e^{\frac{Q_A(t)}{\beta}} + e^{\frac{Q_B(t)}{\beta}}}$$

where β is an inverse gain parameter and reflects the participant's tendency to exploit (stick with responses that have yielded reward in the past) vs explore (try out different choices to determine whether a more rewarding option is available). The model was fit to each subject's trial-by-trial choices by adjusting the three free parameters ($\propto_G$, $\propto_L$,and β) to maximize the log likelihood estimate (LLE):

$$LLE = \log\left(\prod_t P_{i*,t}\right)$$

This was accomplished by incrementing each of the parameters from 0.01 to 1 by 0.03 and selecting the combination of parameters that maximized LLE (i.e. that best predicted the actual choices made by the subject on a trial-by-trial basis). Model fits for each subject were characterized using LLE values as well as Bayesian Information Criterion (BIC) (Schwarz 1978), which penalizes for the number of free parameters used to protect against overfitting.

Model-based fMRI analyses were conducted by including trial-by-trial values of Q and prediction error as parametric regressors in the general linear model. The Q value of the chosen stimulus was included as a regressor at the time of choice, and prediction error was coded at the time of feedback. In the first analysis, positive and negative prediction errors were included in one regressor for signed prediction error; in a second analysis, positive and negative prediction errors were coded separately. As described above, a regressor for time was also included in these GLMs, and regions displaying significant effects of the parametric regressors were identified using whole-brain and ROI-level ANOVA analyses including time as a factor.

Results

Demographic and Clinical Characteristics

Participant demographic and clinical characteristics are shown in Table 1. There were no significant differences between patients and controls in age, sex, race, or handedness. Personal education was higher among controls than patients, an expected finding given the effects of schizophrenia on function, but parental education (a surrogate for premorbid socioeconomic status) was similar between groups. Smoking status also did not differ significantly between groups, both in terms of the number of participants who smoke, and the Fagerstrom nicotine dependence scores among smokers.

SANS scores for positive and negative symptoms were higher among patients than controls, though disorganization scores were low in patients and did not differ between groups. Anhedonia and avolition scores were higher among patients than controls on all clinical and self-report measures except the TEPS anticipatory pleasure score. All but three patients were taking antipsychotic medications, which were primarily atypical antipsychotics, and a number of patients were taking additional medications including antidepressants, mood stabilizers, anxiolytics, and anticholinergics. Two controls and seven patients had a lifetime history of Major Depressive Disorder, and two controls and seven patients had a lifetime history of substance abuse/dependence.

Table 1: Clinical and Demographic Characteristics

	Mean (SD)	
	CON	SCZ
N	37	38
Age	36.43(8.44)	35.00(9.25)
Education (years)	14.14(2.10)	12.95(2.29)*
Highest Parental Education (years)	13.78 (1.64)	18.84 (3.43)
Sex (% Male)	43.2	63.2
Race (% Caucasian)	29.7	42.1
Smoking Status (% Smokers)	43.2	57.9
FagerstromTest for Nicotine Dependence	7.4(5.488)	9.21(4.023)
Handedness	4.31(0.95)	4.97(0.19)
SAPS/SANS Positive	0.03(0.16)	3.5(2.64) *
SAPS/SANS Negative	1.49(2.09)	7.92(2.97) *
SAPS/SANS Disorganization	2.38(1.30)	2.63(2.65)
BNSS Total Anhedonia	0.68(2.15)	1.05(1.91) *
BNSS Total Avolition	5.76(3.48)	5.38(2.44) *
Chapman Social Anhedonia	8.92(6.26)	15.39(7.98) *
Chapman Physical Anhedonia	11.78(6.12)	18.53(10.13) *
Snaith-Hamilton Pleasure Scale	52.38(3.21)	48.45(8.21) *
TEPS Anticipatory Pleasure	48.22(5.69)	46.05(8.31)
TEPS Consummatory Pleasure	38.41(5.64)	35.03(7.37) *
Apathy Scale	22.19(3.57)	26.11(7.26) *
Past Major Depressive Disorder (#)	2	7
Past Substance Dependence (#)	2	7
Antipsychotic Medication (# Taking)	-	35
Typical Only	-	4
Atypical Only	-	29
Typical + Atypical	-	2
Clozapine Only	-	2
Clozapine + Other	-	3
Antipsychotic Dose (CPZ equivalent)	-	717.78(474.25)
Antidepressant Medication (# Taking)	-	17
Antianxiety Medication (# Taking)	-	6
Mood Stabilizer Medication (# Taking)	-	9
Anticholinergic Medication (# Taking)	-	12

*p < .05; CON = control, SCZ = schizophrenia, SD = standard deviation; SAPS/SANS = Scale for the Assessment of Positive/Negative Symptoms; BNSS =Brief Negative Symptom Scale; TEPS = Temporal Experience of Pleasure Scale; CPZ = Chlorpromazine

fMRI Movement Analysis

Ten individuals with schizophrenia and 4 healthy controls were excluded for movement that exceeded a predetermined criterion (mean voxelwise standard deviation values of 20 for 4 or more BOLD runs). In addition, one patient chose to end the experiment early and did not complete the final BOLD run or the test phase. Data for absolute movement, frame-by-frame movement, and mean voxelwise standard deviation for the remaining 38 patients and 37 controls is shown in Table S1. Repeated-measures ANOVAs with BOLD run (1-10) as a within-subjects factor and Group (CON, SCZ) as a between-subjects factor revealed no significant main effects of group or BOLD run X group interactions for any of the three measures (all p values > .27), indicating that movement did not differ between patients and controls. However, there was a main effect of BOLD run in all three measures (absolute: $F(9,648)=2.97, p< .02$; frame-by-frame: $F(9,648)=3.20$, $p<.002$; SD: $F(9,648)=6.42$, $p<.001$), indicating that movement increased significantly over time in both groups.

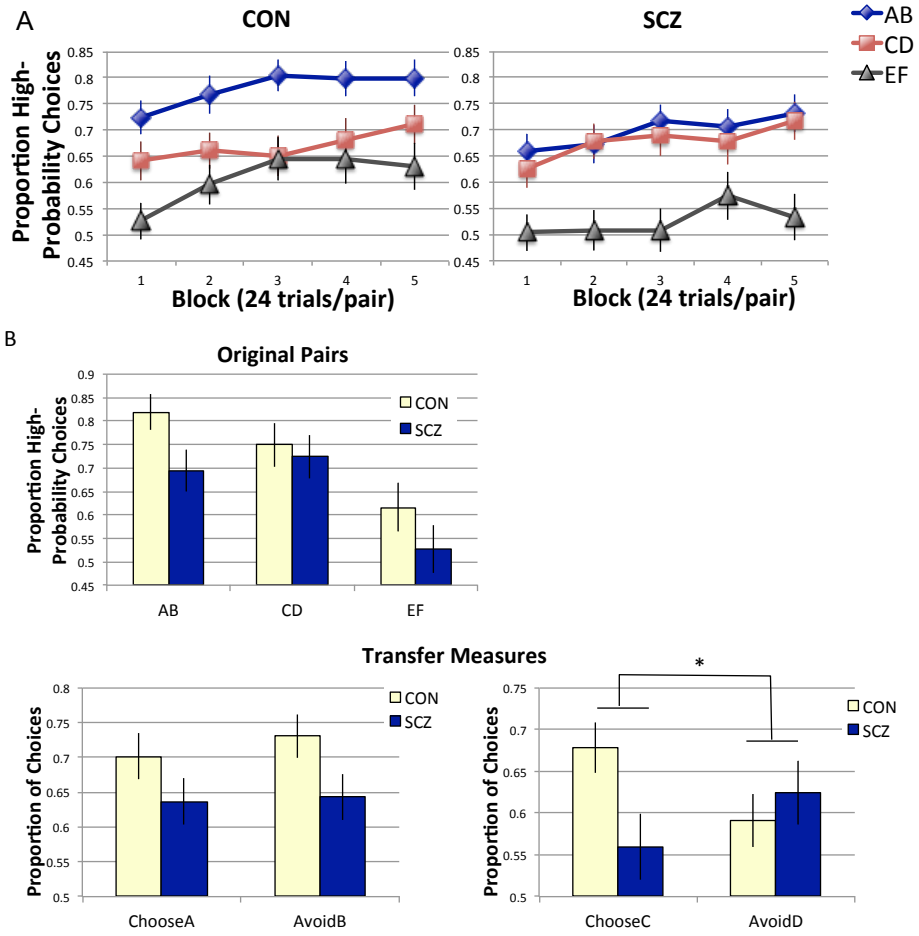
Behavioral Analysis

Acquisition Phase: Choice behavior during the acquisition phase was analyzed by calculating the proportion of trials in which participants chose the more frequently rewarded stimulus (A, C, or E) over the less-frequently rewarded stimulus (B, D, or F). To examine how this proportion changed over the course of learning, the acquisition phase was divided into 5 blocks of 72 trials each (24 trials per stimulus pair), which were used in a repeated measures ANOVA with Block (1-5) and Pair (AB, CD, EF) as within-subjects factors and Group (CON, SCZ) as a between-subjects factor. This analysis revealed significant main effects of Block ($F(4,292) = 8.74$, $p<.001$), with increasing proportions of high-probability choices over time, and Pair ($F(2,146) = 22.00$, $p<.001$), with the greatest proportion of high-probability choices for AB pairs, followed by CD, then EF. The main effect of group was not significant, nor were any interactions (p values > .1). However, post-hoc simple effects tests revealed trend-level group differences within the AB ($F(1,73) = 3.45$, $p < .07$ and EF ($F(1,73) = 3.22$, $p<.08$) pairs, but not the CD pair ($p>.8$). As shown in Figure 3a, performance for the CD pair is similar between patients and controls, whereas patients performed more poorly than controls on the AB and EF pairs.

Test Phase: Test phase choice data is shown in Figure 3b. For the original AB, CD, and EF pairs that had been presented during acquisition, a Pair X Group ANOVA revealed a main effect of pair

($F(2,144) = 13.81$, $p < .001$), but no main effect of group ($F(1,72) = 2.67$, $p > .1$) or Pair X Group interaction ($F(2,144) = 0.86$, $p > .42$). However, post-hoc simple effects tests revealed a significant group difference for the AB pair only ($F(1,72) = 4.44$, $p < .04$), in which patients performed more poorly than controls. Performance on the CD and EF pairs did not differ significantly between groups. In order to examine whether learning from positive versus negative feedback differed between groups, we compared ChooseA and AvoidB measures derived from test phase performance. The ChooseA measure was created by averaging performance on test phase trials in which stimulus A was presented in a novel pairing against a less frequently reinforced stimulus (AC, AD, AE, AF). Similarly, the AvoidB measure (BC, BD, BE, BF) was created as an index of learning from negative outcomes. Performance on these transfer measures was compared using a repeated measures ANOVA with ChooseA vs AvoidB as a within-subjects factor and Group as a between subjects factor, which revealed a trend-level main effect of Group ($F(1.72) = 3.61$, $p <.07$), but no main effect of or interaction with the ChooseA vs Avoid B factor. As shown in Figure 3b, individuals with schizophrenia performed more poorly than controls on both the ChooseA and AvoidB measures. However, it is important to note that while the ChooseA/AvoidB measure has been the transfer measure of interest in previously published versions of this task, the appropriateness of this measure in this sample is called into question by the fact that patients performed more poorly on the AB pair than controls. This raises the possibility that the sensitivity of transfer measures relying on this pair to pick up any bias toward choosing A or avoiding B is reduced in the patient group. To avoid this problem, an equivalent transfer measure was created that relies on the CD pair, performance on which was very closely matched between groups: ChooseC (CE, CF) vs. AvoidD (DE, DF). ANOVA analysis of this measure revealed a significant ChooseC/AvoidD X Group interaction ($F(1,72) = 5.21$, $p< .03$), with no significant main effects (p values > .2). As shown in Figure 3b, ChooseC performance was significantly lower in patients than controls ($t(72) = 2.40$, $p <.02$), while performance on the AvoidD measure did not differ significantly between groups ($p>.5$).

Figure 3: Behavioral results. (A) Acquisition phase performance. Proportion of high-probability choices (A/C/E) per 24-trial block. (B) Test-phase performance for original AB, CD, and EF pairs, and ChooseA/AvoidB and ChooseC/AvoidD transfer measures.



Modeling Results: As an additional measure of positive versus negative learning, a Q-learning algorithm with separate estimates of positive and negative learning rate was fit to the behavioral data for each subject. Model fits as demonstrated by LLH and BIC values (in which smaller numbers indicate better fit) did not differ significantly between groups, as shown in Table 2. However, there were a number of subjects in both groups who showed no appreciable learning during the task. Model fits for these subjects were poor, in many cases demonstrating BIC values that exceeded the null BIC associated with chance. In order to restrict the modeling analysis to those subjects whose choices were well-described by the model, a subset of "nonlearners" was excluded. This group was defined as those participants who failed to perform significantly above chance on the AB pair during the second half of the acquisition

phase, and consisted of 11 patients and 10 controls. Model fits remained similar between patients and controls after exclusion of nonlearners (Table 2).

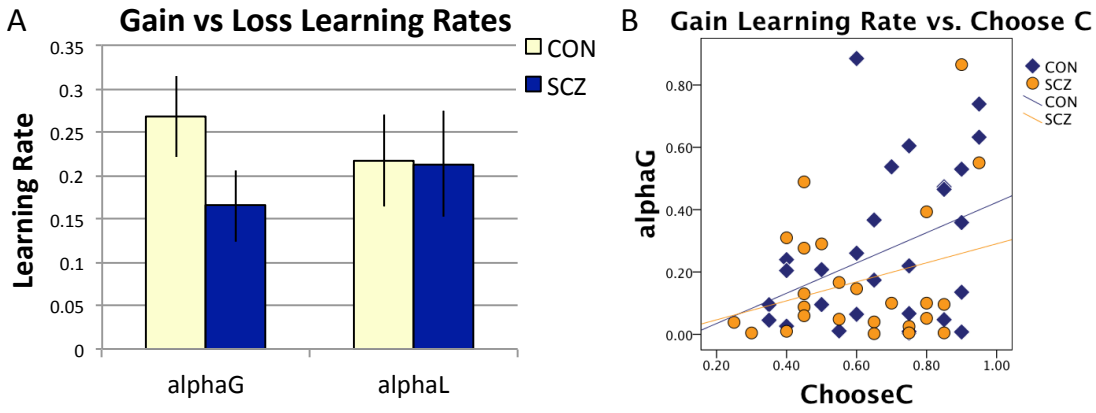
Table 2: Model Fit Data

Sample	Criterion	CON	SCZ
Full Sample	LLH	-173.13 (57.56)	-184.68 (52.90)
(38 CON, 37 SCZ)	BIC	363.83 (115.13)	386.89 (105.85)
Learners Only	LLH	-154.6 (52.98)	-171.62 (53.78)
(28 CON, 26 SCZ)	BIC	326.78 (105.96)	360.83 (107.55)

LLH = Log likelihood; BIC = Bayesian Information Criterion

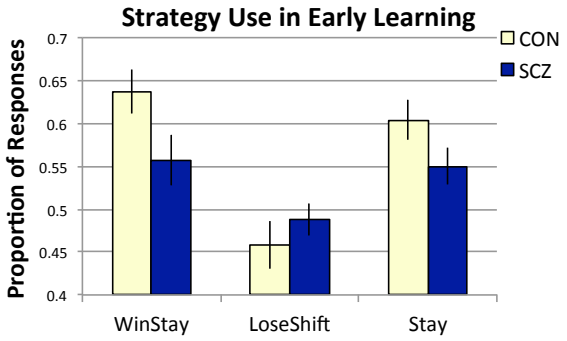
To compare learning rates from positive and negative feedback between groups, fitted learning rate parameters were compared in a Learning Rate (Gain, Loss) X Group ANOVA. As shown in figure 4a, the pattern seen for gain and loss learning rates is similar to that seen for the ChooseC and AvoidD transfer measures, which index learning from gain and loss respectively. However, the learning rate ANOVA failed to reveal a significant learning rate X group interaction ($F(1,52) = 1.24$, $p>.2$) or significant main effects of learning rate or group (p values > .3). Of note, there were 3 outliers in the patient group, the removal of which yielded a trend-level main effect of group ($F(1,49) = 3.15$, $p<.09$) and a post-hoc group difference in Gain learning rate ($t(49) = 2.09$, $p<.05$). Given that gain learning rate and the ChooseC transfer measure both index a tendency to learn from positive outcomes, we expected a positive correlation between these two measures; analogously, we also expected a positive relationship between loss learning rate and AvoidD because both of these measures index learning from negative outcomes. Partial correlations controlling for group revealed a significant positive relationship between ChooseC and gain learning rate ($r(50)=.333$, $p<.02$; see figure 3b), but not between AvoidD and loss learning rate ($p>.9$). To test the specificity of the relationship between gain learning rate and ChooseC, we also correlated ChooseC with loss learning rate and AvoidD with gain learning rate. Neither correlation was significant (p values > .4).

Figure 4: Model parameter comparisons. (A) Gain and loss learning rates for patients and controls. (B) Correlations between gain learning rate and ChooseC transfer measure among patients and controls.



Strategy Analysis: Previous studies of reinforcement learning in schizophrenia have revealed that patients often rely on explicit strategies during early learning to a lesser extent than controls (Gold et al 2008). To examine this possibility, we examined use of Win-Stay/Lose-Shift strategy among patients and controls during the first 4 acquisition runs (48 trials per pair). This was accomplished by calculating the proportion of Wins (positive feedback) that were followed by Stays (choosing the same stimulus again the next time the same pair was presented) and the percentage of losses (negative feedback) that were followed by shifts (choosing the opposite stimulus the next time the pair was presented). Independent samples t-tests revealed a significant difference in win-stay proportion ($t(73) = 2.08$, $p<.05$), in which patients showed a lower proportion of stays after wins than controls (Figure 5). Lose-shift proportion did not differ between groups ($p>.3$). In order to examine whether the difference in win-shift activity was related to a more general tendency among patients to shift more often regardless of feedback, we also examined the total proportion of stays versus shifts. An independent samples t test was significant at a trend level ($t(73)=1.71$, $p<.1$), with patients staying less often (and therefore shifting more often) overall than controls. Similar results were obtained when the analysis was restricted to the first 2 acquisition runs.

Figure 5: Win-Stay/Lose-Shift strategy use during the first four learning blocks in patients and controls. WinStay: the proportion of wins that were followed by stays; LoseShift: the proportion of losses that were followed by shifts; Stay: the total percentage of stays.



Individual Differences Analyses: We had hypothesized that clinical symptoms of anhedonia/amotivation would be related to a reduced ability to learn from positive outcomes in individuals with schizophrenia. To examine this possibility, we conducted correlations between the summary anhedonia/amotivation scores created from our clinical and questionnaire measures and ChooseC scores among all patients. We also correlated anhedonia/amotivation scores with gain learning rates among learners. Neither analysis revealed a significant relationship between anhedonia score and Go learning (p values > .1). Finally, we correlated win-stay rates with anhedonia/avolition scores, which also failed to reveal a significant relationship (p > .1). Because several studies in the literature reported correlations with total negative symptoms rather than with anhedonia/avolition specifically, we also conducted the above correlations with SANS negative symptom scores, but failed to find significant relationships with this measure as well.

Behavior Summary: In summary, our analyses of choice behavior during both acquisition and test are suggestive that patients chose A over B less often than controls, although these effects were small. On transfer measures designed to assess learning from positive vs. negative feedback, patients demonstrated a specific reduction in learning from positive feedback. Gain vs. loss learning rates derived from a reinforcement learning model showed a similar (albeit weaker) pattern, and there was a significant positive relationship between the model- and transfer-derived measures of learning from positive feedback in both groups. Early in learning, patients demonstrated a greater tendency to shift after wins than controls. Contrary to our predictions, none of the behavioral measures of interested showed a significant relationship with anhedonia/avolition severity.

fMRI Analysis

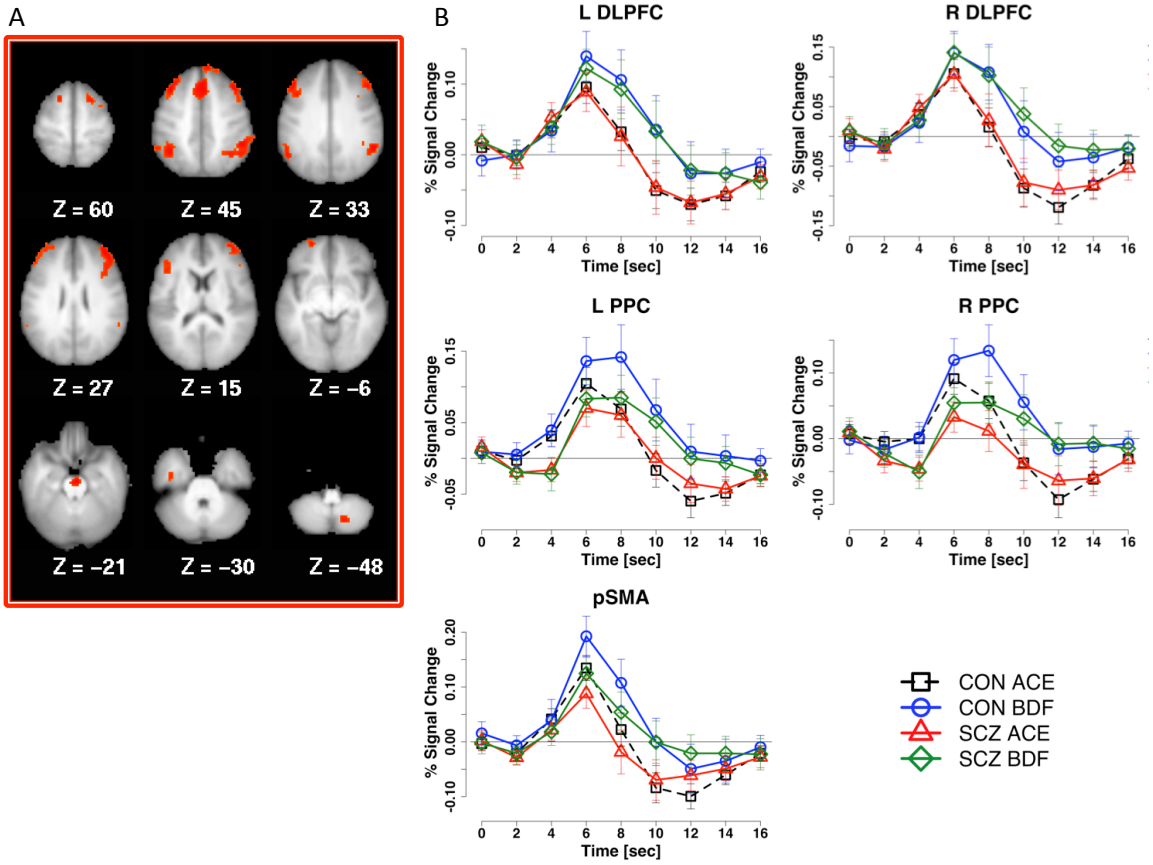
Choice-related activation

Analysis 1: Full acquisition phase; collapsed across trial type. GLMs for the first analysis included regressors for two choice types (A/C/E, B/D/F). Stimulus pairs were combined for the purposes of this analysis because a number of subjects had too few "B" choices to model activity for each pair type separately. Effects that differed across stimulus pair were addressed in a separate analysis that excluded these subjects, described below. In analysis 1, we examined activity associated with choices of high vs. low-probability stimuli, using a repeated-measures ANOVA with Choice (A/C/E, B/D/F) and Time (1-9) as within-subjects factors and Group (CON, SCZ) as a between-subjects factor.

A priori ROI analysis within bilateral caudate, putamen, and nucleus accumbens revealed significant main effects of time in all regions, but no significant interactions with time. Whole-brain analysis revealed a number of regions demonstrating a significant choice x time interaction, shown in Figure 6a and Table 3. For clarity, only regions demonstrating a significant choice x time effect within both groups separately on post-hoc tests are included here. This analysis revealed a significant choice x time effect in several members of the cognitive control network (Cole and Schneider 2007) including bilateral DLPFC, posterior parietal cortex, cerebellum, and ACC/preSMA. All of these regions demonstrated greater activation for low-reward than high-reward choices for both patients and controls (Figure 6b). In addition, a region in dorsomedial prefrontal cortex demonstrated greater deactivation for high-probability than low-probability choices in both groups. Table 3 also includes post-hoc tests conducted within each stimulus pair, and shows that most of these patterns were driven by the AB and/or EF pairs, with little contribution of the CD pair.

Figure 6: Choice ANOVA results: main effect of choice in the full acquisition phase. (A) Regions demonstrating a main effect of choice on whole-brain analysis ($Z \geq 3$, $n \geq 13$). All regions showed greater activity for incorrect than correct choices. (B) Example timecourses for cognitive control regions,

demonstrating greater activation for incorrect than correct choices in both patients and controls.



Time x Group interactions (Table 3) were seen in regions in bilateral superior parietal lobule, left superior frontal gyrus, and cuneus, which demonstrated greater activation overall (i.e. for both choice types) among controls than patients. Greater activation for patients than controls was seen in bilateral cuneus, left inferior parietal lobule, and right superior parietal lobule.

Table 3. Choice ANOVA: Regions demonstrating main effects

Effect		Region	BA	Coordinates	# Voxels	Z	Pattern	AB	CD	EF
Choice x Time	R	Cerebellar Tonsil	-	+16, -66, -49	19	4.32	BDF>ACE	NS	NS	***
	L	Middle Frontal Gyrus	9	-40, +27, +33	267	5.12	BDF>ACE	NS	NS	***
	R	Middle Frontal Gyrus	9	+39, +28, +34	366	6.61	BDF>ACE	***	NS	***
	R	Pons	-	+2, -19, -20	23	5.09	BDF>ACE	NS	NS	***
	L	Middle Frontal Gyrus	10	-26, +54, -7	15	4.62	BDF>ACE	NS	NS	***
	R	Inferior Parietal Lobule	40	+41, -53, +44	285	6.21	BDF>ACE	***	NS	***
	L	Inferior Parietal Lobule	40	-45, -51, +41	236	5.00	BDF>ACE	***	NS	***
		Medial Frontal Gyrus	8	+0, +21, +46	228	4.71	BDF>ACE	***	NS	**
	L	Middle Frontal Gyrus	6	-19, +11, +62	18	4.96	BDF>ACE	NS	NS	***
	R	Superior Frontal Gyrus	6	+18, +11, +63	44	5.84	BDF>ACE	*	NS	***
	R	Medial Frontal Gyrus	8	+6, +49, +42	33	4.43	(-) ACE>BDF	NS	NS	***
	L	Uncus	20	-38, -15, -28	23	4.31	ACE: early peak, then (-)	NS	NS	***
Time x Group	R	Precuneus	7	+27, -66, +30	16	3.87	CON > SCZ	***	***	**
	L	Superior Frontal Gyrus	10	-18, +52, -3	42	4.95	CON > SCZ	***	**	***
	L	Superior Frontal Gyrus	6	-1, +19, +58	94	3.37	CON > SCZ	*	NS	***
	R	Superior Parietal Lobule	7	+15, -67, +59	63	4.43	CON > SCZ	***	***	***
	R	Cuneus	19	+14, -91, +24	43	4.14	SCZ > CON	***	***	**
	L	Cuneus	19	-21, -91, +26	34	4.71	SCZ > CON	***	***	***
	L	Inferior Parietal Lobule	40	-52, -32, +54	14	3.65	SCZ > CON	***	***	**
	R	Superior Parietal Lobule	7	+32, -45, +61	15	3.13	SCZ > CON	***	***	*
	L	Insula	13	-44, -19, +0	13	3.50	(-) CON > SCZ	NS	***	**
	R	Inferior Temporal Gyrus	20	+55, -23, -18	18	3.38	(+) SCZ, (-)CON	***	NS	*

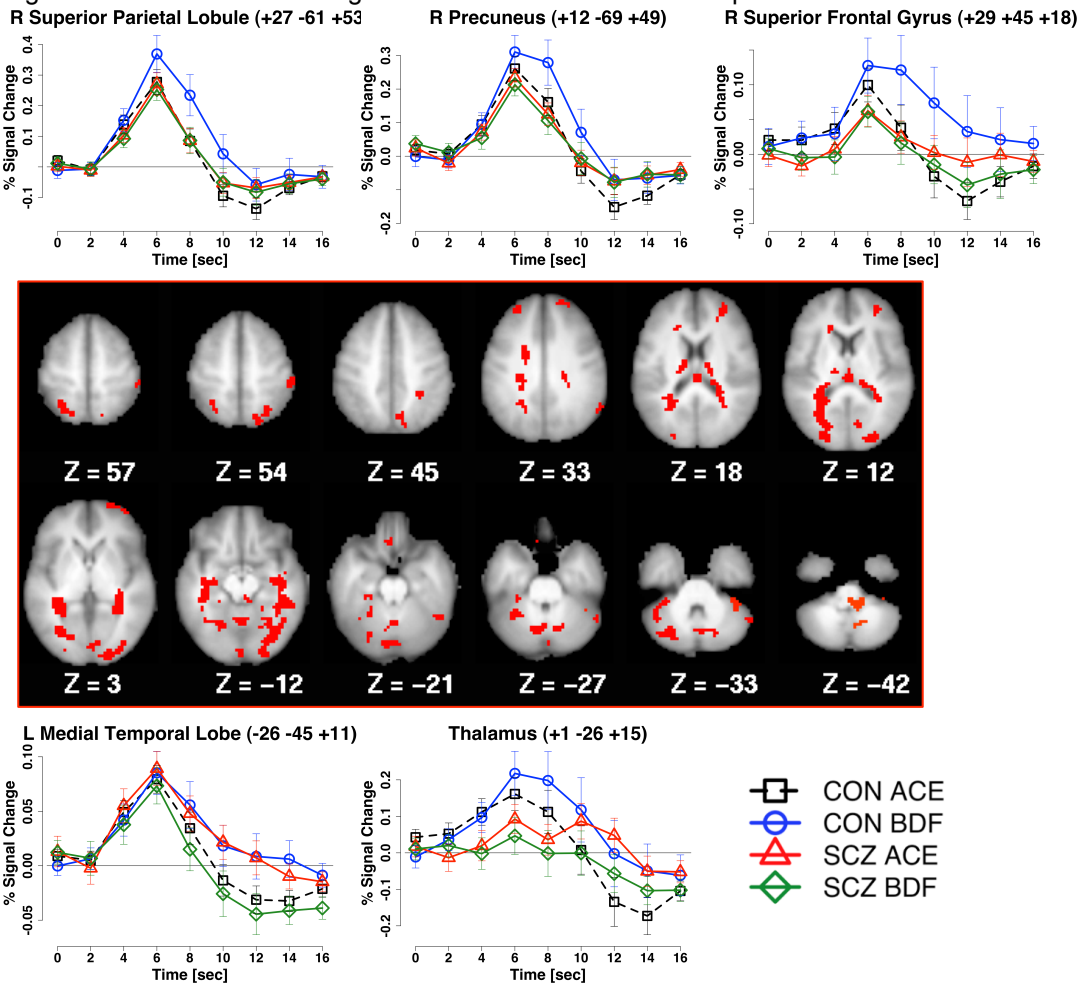
Regions showing a significant Choice x Time x Group interaction are shown in Figure 7 and Table 4. The majority of these regions, those predominantly located in the cerebellum and occipital lobe, activated more strongly for low- than high-probability choices in controls, but for high- than low-probability choices in patients. Another subset of regions including bilateral precuneus also showed greater activation for low-probability choices in controls, but no differentiation between choice types in patients. Finally, regions including right superior frontal gyrus showed greater activation for high- than low-probability choices in patients, and no differentiation among controls.

Table 4. Choice ANOVA: regions demonstrating Choice x Time x Group interactions

	Region	BA	Talairach	Voxels	Z	Post-hoc	
						CON	SCZ
L	Cerebellar VI	-	-31_-46_-26	112	6.10	BDF > ACE	ACE > BDF
L	Fusiform Gyrus	-	-24_-62_-13	42	4.00	BDF > ACE	ACE > BDF
R	Cerebellar Vermis VI	-	+3_-65_-24	134	4.28	BDF > ACE	ACE > BDF
L	Crus I	-	-29_-72_-33	34	3.88	BDF > ACE	ACE > BDF
R	Cerebellar VI	-	+40_-43_-34	57	5.11	BDF > ACE	ACE > BDF
R	Cerebellar IX	-	+10_-51_-39	96	5.09	BDF > ACE	ACE > BDF
R	Fusiform Gyrus	19	+27_-74_-12	270	6.92	BDF > ACE	ACE > BDF
R	Inferior Parietal Lobule	40	+52_-31_+54	17	4.52	BDF > ACE	ACE > BDF
L	Middle Occipital Gyrus	18	+24_-90_+6	98	4.00	BDF > ACE	ACE > BDF
L	Precentral Gyrus	4	-33_-27_+67	34	3.87	BDF > ACE	ACE > BDF
L	Rectal Gyrus	11	-3_+23_-22	16	4.16	BDF > ACE	ACE > BDF
R	Superior Frontal Gyrus	9	+19_+51_+31	16	3.59	BDF > ACE	ACE > BDF
R	MTL/Occipital WM	-	+32_-35_+6	462	6.63	BDF > ACE	ACE > BDF
L	MTL/Occipital WM	-	-26_-45_+11	801	6.17	BDF > ACE	ACE > BDF
L	Superior Parietal Lobule	7	-25_-59_+56	30	4.30	BDF > ACE	ACE > BDF
R	Thalamus	-	+1_-26_+15	37	4.39	BDF > ACE	ACE > BDF
L	Cerebellar Crus I	-	-26_-75_-22	13	2.99	BDF > ACE	NS
R	Cerebellar Vermis VI	-	-1_-85_-22	31	4.25	BDF > ACE	NS
R	Cingulate Gyrus	23	-2_-21_+28	24	4.48	BDF > ACE	NS
R	Cuneus	23	+10_-75_+10	17	2.91	BDF > ACE	NS
R	Precuneus	7	+28_-43_+44	15	3.74	BDF > ACE	NS
L	Precuneus	7	+12_-69_+49	49	3.95	BDF > ACE	NS
L	Superior Frontal Gyrus	10	+29_+45_+18	55	4.19	BDF > ACE	NS
L	Superior Frontal Gyrus	9	-27_+45_+32	14	3.86	BDF > ACE	NS
R	Superior Parietal Lobule	7	+27_-61_+53	14	3.90	BDF > ACE	NS
L	Cerebellar Peduncle	-	-19_-56_-40	42	4.94	ACE > BDF early	ACE > BDF
R	Cingulate WM	-	+19_+0_+32	41	3.85	ACE > BDF early	ACE > BDF
R	Cerebellar Culmen	-	+8_-37_-13	16	3.42	NS	ACE > BDF
R	Superior Frontal Gyrus	10	+29_+59_+2	24	4.35	NS	ACE > BDF
L	Frontal WM	-	-18_+25_+16	19	3.62	NS	ACE > BDF
R	Angular Gyrus	39	+57_-59_+34	14	4.86	NS	(-) ACE > BDF

R = Right, L = Left, WM = White Matter, BA = Brodmann Area, MTL = Medial Temporal Lobe

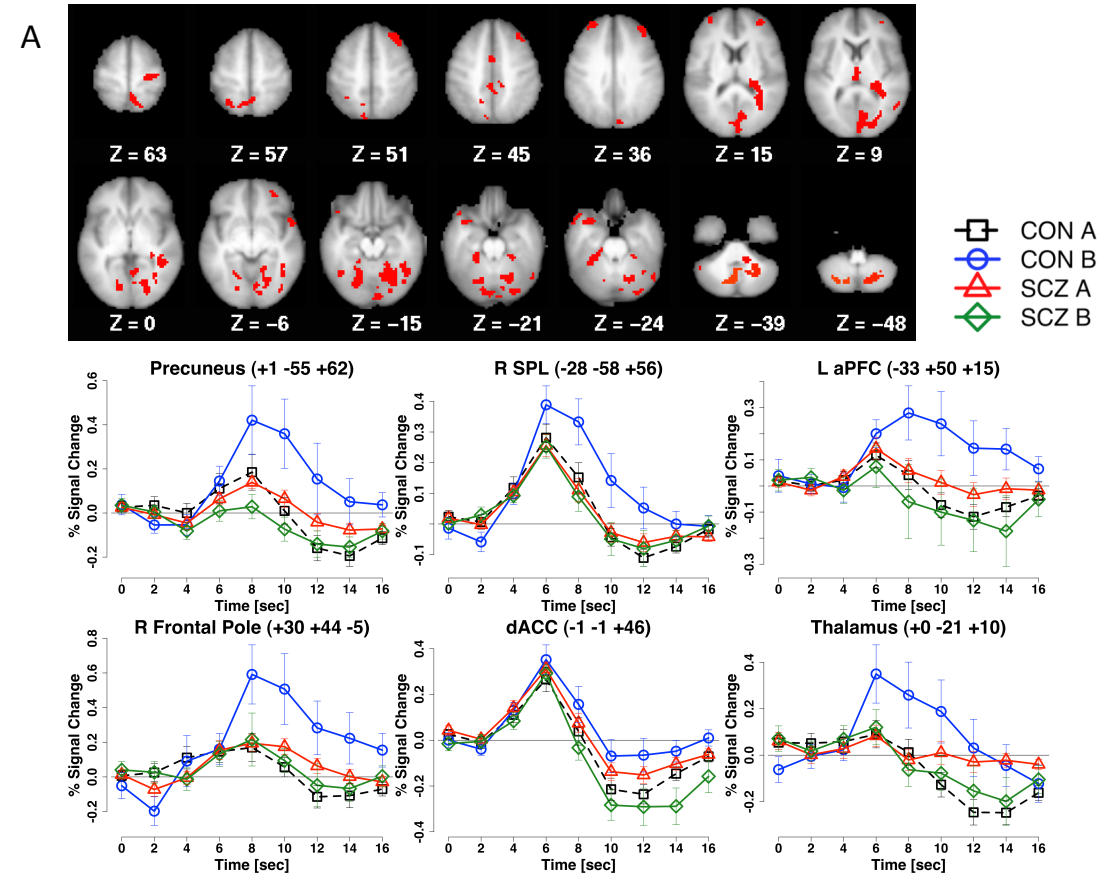
Figure 7. Choice ANOVA: Regions with Choice x Time x Group interactions



ANOVA Analysis 2: within AB pair. Given that the greatest group differences in behavior were seen for the AB pair, we conducted an additional analysis to examine choice effects within this pair. After exclusion of subjects who failed to make sufficient numbers of B choices, which yielded an N of 32 controls and 38 patients, a Choice (A,B) x Time (1-9) x Group (CON, SCZ) ANOVA was conducted to examine choice-related activity for the AB pair. The results of this analysis showed considerable overlap with the results of the previous analysis. The striatal ROI analysis again revealed no significant effects or interactions with choice or group. On whole-brain analysis, the Choice x Time results were nearly identical, with several cognitive control regions demonstrating greater activation for "B" choices than "A" choices among both patients and controls. The Choice X Time X Group results were also similar, with a number of cerebellar and occipital regions showing significant interactions. However, within the AB pair,

additional regions were with greater responses to “B” choices among controls, but not patients, were identified. These included regions typically associated with error or conflict processing (Dosenbach et al 2006), such as bilateral anterior prefrontal cortex, dorsal ACC, and thalamus (Figure 8a). To examine how this activation related to performance, we conducted partial correlations controlling for group between activation for B choices in these regions and average performance on the AB pair. Greater activation was associated with better AB performance among both patients and controls in several of the regions including left aPFC ($r = .256$, $p < .04$) and left superior parietal lobule ($r = .452$, $p < .001$), with several more regions including dACC and right aPFC showing trend-level correlations (Figure 8b). These correlations were driven more by controls than patients, but were in the same direction in both groups. Left superior parietal lobule also correlated with win-stay (but not lose-shift) behavior for the AB pair among patients ($r = .446$, $p < .007$) and controls ($r = .518$, $p < .004$).

Figure 8. Choice ANOVA: within AB pair. (A) Regions and example timecourses for significant Choice x Time x Group interactions. (B) Correlations between activation for B choices and performance on the AB pair (proportion of "A" choices during the full acquisition phase), or early Win-Stay performance (proportion of wins followed by stays during the first 2 blocks).



In sum, analyses 1 and 2 revealed no group differences in striatal activation for choices associated with higher versus lower probability of reward. This was true for all stimulus pairs combined, and within the AB pair. In addition, both groups showed greater activation in many cognitive control regions for lower-probability than higher-probability choices. However, group differences were seen in a

number of regions, including regions implicated in error or conflict processing (particularly for the AB pair), which activated more strongly for low-probability choices among controls, but not patients. In one of these regions, left superior parietal lobule, activation for B choices was also associated with higher proportions of A choices and more frequent win-stay behavior in both groups.

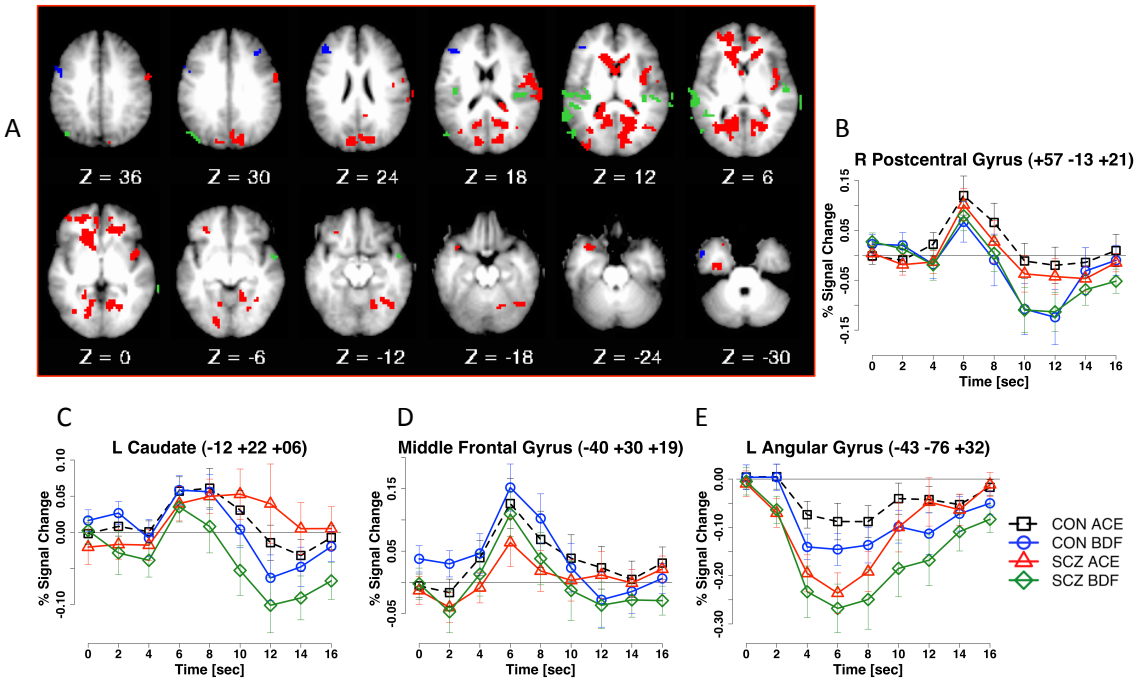
Analysis 3: Learning Phase. The previous analyses examined choice-related activation throughout the full course of the acquisition phase, and therefore did not take into account any changes in this activity that may have occurred during the course of learning. In order to examine such changes, the acquisition phase was divided into Early and Late learning phases comprising the first and last 60 trials per stimulus pair, respectively. As shown in Figure 3a, group-level performance began to plateau during Block 3, suggesting that on average the Early learning phase represents a period of ongoing learning, while the Late phase represents the learning plateau. ANOVAs for choice-related activity collapsed across all pairs and within the AB pair alone, analogous to analyses 1 and 2 above, were conducted within the early learning phase to examine activation patterns that may have been present during learning, but obscured by post-plateau trials in the previous analyses. In addition, a repeated measures ANOVA using Learning Phase (Early, Late), Choice, and Time as within-subjects factors and Group as a between-subjects factor was conducted to identify activation patterns that differed across learning stage.

Table5: Choice ANOVA results within early learning phase

Effect		Region	BA	Talairach	Voxels	Z	Pattern		Post-Hoc	
							CON	SCZ	CON	SCZ
Choice x Time	R	Cuneus	18	+11_-70_+13	523	5.02	ACE > BDF		*	*
	L	Inferior Frontal Gyrus	47	-29_+28_-8	14	4.28	ACE > BDF		*	*
	L	Inferior Frontal Gyrus	10	-36_+43_+0	41	4.90	ACE > BDF		*	*
	R	Anterior Insula	13	+39_-3_+7	104	4.63	ACE > BDF		*	*
	R	Middle Occipital Gyrus	19	+40_-66_+13	15	4.04	ACE > BDF		*	*
	R	Postcentral Gyrus	3	+57_-13_+21	137	6.37	ACE > BDF		*	*
	L	Posterior Cingulate	30	-17_-62_+7	299	5.23	ACE > BDF		*	*
	L	Caudate	-	-12_+22_+06	155	5.27	ACE > BDF		NS	*
	R	Caudate	-	+16_+27_+03	99	4.64	ACE > BDF		NS	*
	L	Cerebellar Peduncle	-	-26_-51_-42	20	4.70	ACE > BDF		NS	*
	R	Fusiform Gyrus	37	+37_-60_-13	47	4.83	ACE > BDF		NS	*
	L	Middle Occipital Gyrus	19	-35_-88_+11	14	4.11	ACE > BDF		NS	*
	R	Middle Frontal Gyrus	9	+35_+26_+28	14	3.63	ACE > BDF		NS	*
	L	Middle Frontal Gyrus	46	-40_+30_+19	44	4.56	BDF > ACE		*	*
	L	Precentral Gyrus	6	-52_+4_+34	18	4.31	BDF > ACE		*	*
	L	Angular Gyrus	39	-43_-76_+32	25	4.52	(-) BDF > ACE		*	*
	R	Posterior Insula	13	+42_-28_+15	44	4.63	(-) BDF > ACE		*	*
	L	Middle Temporal Gyrus	39	-51_-73_+13	30	4.82	(-) BDF > ACE		*	*
	R	Superior Temporal Gyrus	21	+53_-3_-9	17	4.65	(-) BDF > ACE		*	*
	R	Superior Temporal Gyrus	22	+70_-36_+7	26	5.93	(-) BDF > ACE		*	*
	L	Superior Temporal Gyrus	22	-54_-37_+10	184	5.99	(-) BDF > ACE		*	*
Time x Group	L	Thalamus	-	-1_-31_+11	43	3.36	CON > SCZ		-	-
	L	Precuneus	7	-25_-67_+36	62	4.82	CON > SCZ		-	-
	L	Superior Parietal Lobule	7	-29_-59_+60	27	3.45	CON > SCZ		-	-
	L	Inferior Parietal Lobule	40	-47_-43_+42	43	4.30	CON > SCZ		-	-
	R	Superior Frontal Gyrus	6	+0_+19_+57	169	5.01	CON > SCZ		-	-
	R	Superior Parietal Lobule	7	+25_-62_+48	345	5.67	CON > SCZ		-	-
	R	Precuneus	7	+27_-68_+37	16	3.88	CON > SCZ		-	-
	R	Cerebellar VIIIa	-	+32_-43_-51	15	3.66	CON > SCZ		-	-
	R	Middle Frontal Gyrus	9	+38_+22_+27	21	3.90	CON > SCZ		-	-
	R	Inferior Frontal Gyrus	45	+53_+16_+4	28	4.92	CON > SCZ		-	-
Choice x Time x Group	R	Cerebellar Crus I	-	+55_-61_-30	14	5.54	BDF > ACE	ACE > BDF	*	*
	R	Midbrain	-	+1_-32_-11	14	3.78	BDF > ACE	ACE > BDF	*	*
	R	Thalamus	-	+16_-19_+19	29	4.27	-	ACE > BDF	NS	*
	R	Precentral Gyrus	4	+55_-14_+42	22	4.18	ACE > BDF	-	*	NS

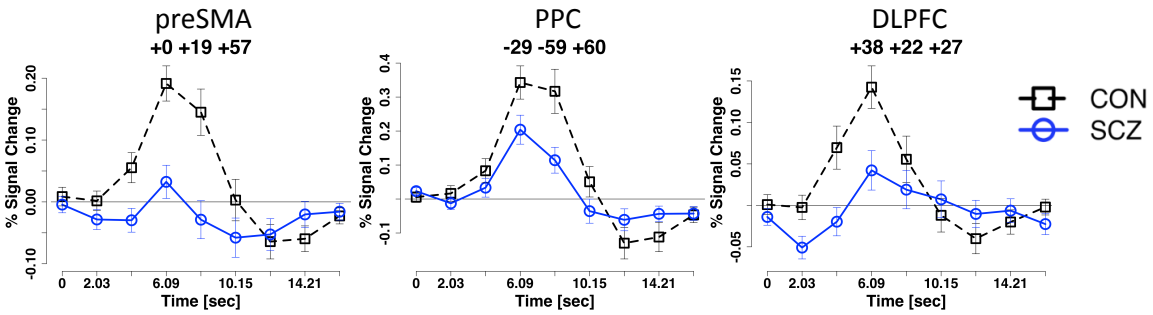
On ROI analysis, a Choice X Time X Group ANOVA within in the early phase once again revealed significant main effects of time in all regions, but no effects of choice or group that survived correction for multiple comparisons. For whole-brain activity, the patterns of activation for high- versus low-probability choices within the early learning phase were quite different than those seen in the full acquisition phase (Table 5). Several regions demonstrated Choice X Time interactions in which activity was greater for high-probability than low-probability choices, including regions in bilateral caudate, left inferior frontal gyrus, right anterior insula, and right postcentral gyrus (Figure 9a). An example timecourse is shown in figure 9b. In a set of regions including bilateral caudate, post-hoc tests within each group separately revealed that the ACE > BDF patterns were significant among patients, but not controls. For example, as illustrated in Figure 9c, controls activated left ventral striatum similarly for both choice types, while in patients, activation was sustained longer for high- than low-probability choices. A pattern of greater activation for low- than high-probability choices was seen in left DLPFC and precentral gyrus, similar to the pattern seen for the full acquisition phase (Figure 9a, blue). As shown in Figure 9d, while activation was lower in patients in these regions overall, both groups showed the same pattern of greater activation for low- than high-probability choices. Lastly, there was a subset of regions that deactivated more strongly for low- than high-probability choices, which included regions in bilateral superior temporal gyrus, right posterior insula, and left angular gyrus (figure 9e).

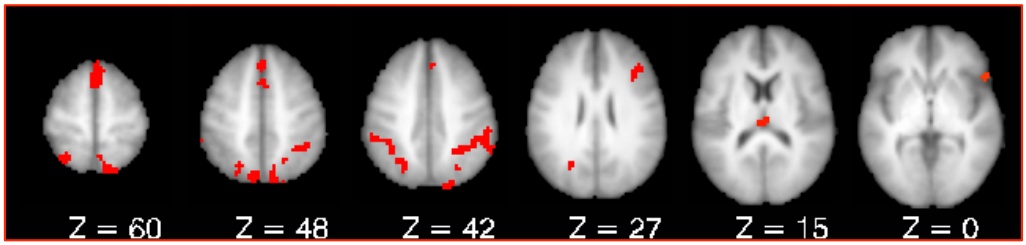
Figure 9. Choice ANOVA during early learning phase: Choice x Time interactions. (A) Regions demonstrating significant Choice x Time interactions. Red = ACE>BDF, Blue = BDF > ACE, Green = deactivation (BDF > ACE).



A Time x Group interaction was present in a number of cognitive control regions including bilateral posterior parietal cortex, right DLPFC, preSMA, thalamus, and right anterior insula/inferior frontal gyrus (Table 5 and Figure 10). All of these regions activated more strongly in controls than patients during the early learning phase regardless of choice type. This pattern is similar to that seen during the full acquisition phase, where preSMA and superior parietal lobule showed greater activity in controls overall; however, in early learning this pattern is present more extensively in regions associated with cognitive control. In addition, while the full acquisition phase also revealed superior parietal regions that activated more strongly in patients than controls, during early learning all significant regions activated more strongly in controls than patients. Choice x Time x Group interactions were also seen in a few regions; among these were midbrain and right cerebellar crus I, which activated more strongly for low than high-probability choices among controls, with the opposite pattern among patients (Table 5).

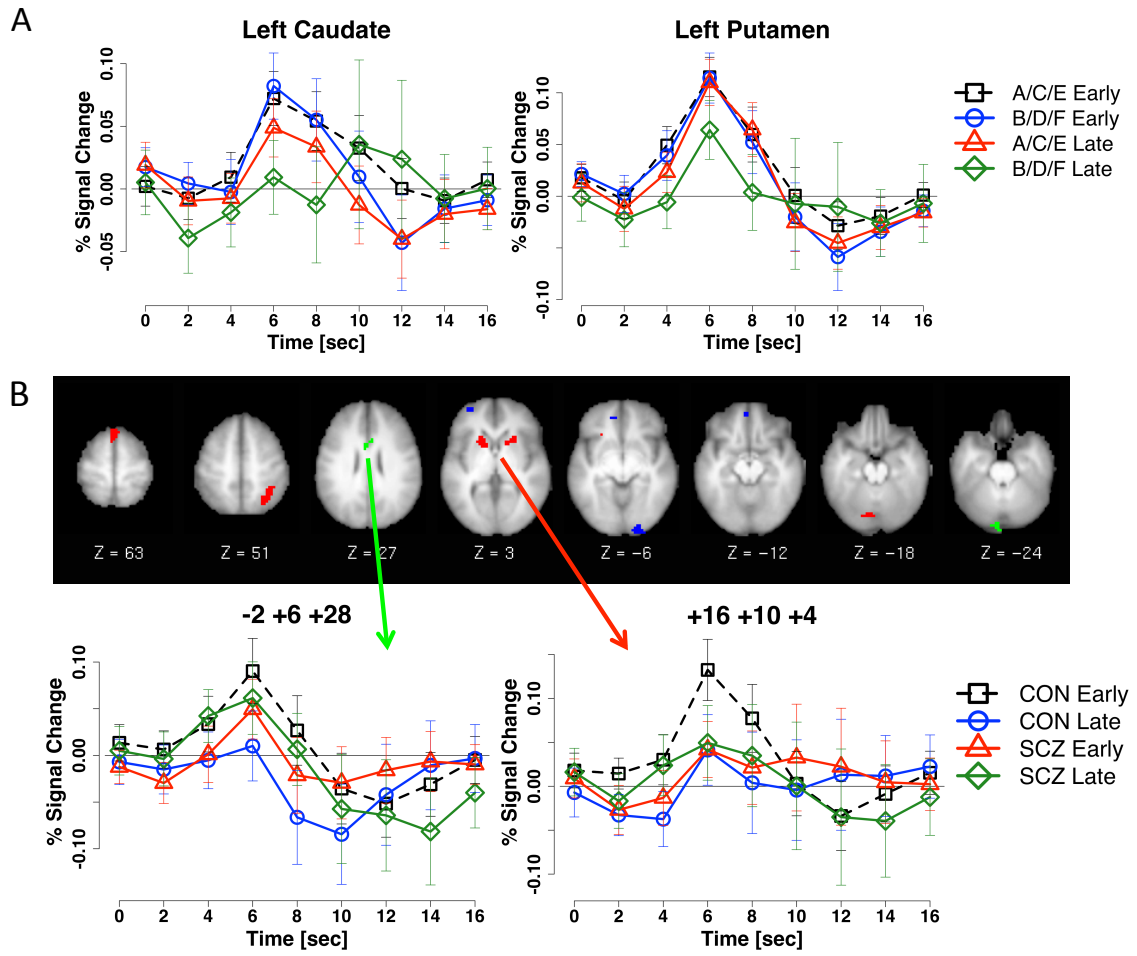
Figure 10. Time x Group interactions in the early learning phase.





We also conducted analyses using learning phase as a factor in order to identify regions whose activity differed between the early and late learning phases. Because some subjects made very few low-probability choices during the late learning phase, a number of subjects were excluded from this analysis, yielding a sample of 32 controls and 33 patients. ROI analysis results revealed significant Learning Phase x Choice x Time interactions in bilateral caudate (L: $F(8,504)=3.59$, $p<.0001$; R: ($F(8,504) = 2.90$, $p<.004$) and bilateral putamen (L: $F(8,504)=2.33$, $p<.02$; R: $F(8,504) = 2.91$, $p<.004$). These regions activated equally for both choice types during the early learning phase, but more strongly for high- than low-probability choices during the late learning phase (Figure 11).

Figure 11. Learning Phase ANOVA results. (A) ROI analysis: example timecourses for regions showing Learning Phase x Choice x Time interactions. (B) Whole-brain analysis: Learning Phase x Time x Group interactions. Red = Early > Late in CON but not SCZ; Blue = Early > Late in SCZ but not CON; Green = Early > Late in CON, Late > Early in SCZ.



On whole-brain analysis, several regions demonstrated a Learning Phase x Time x Group interaction, such that overall activation (irrespective of choice type) changed between early and late learning in a way that differed between groups. These regions are shown in Figure 11b and Table 6. Controls activated regions including bilateral caudate, superior frontal gyrus, and left fusiform gyrus more strongly during the early stages of learning, while patients showed less differentiation between early and late. To determine whether activation in the caudate regions was related to learning, correlations were conducted between early learning phase activation in these regions and learning rate, but did not yield significant relationships.

Table 6. Learning Phase x Time x Group interactions

	Region	BA	Talairach	# Voxels	Z	Pattern	
						CON	SCZ
L	Cerebellar Crus II	-	-6_-95_-24	19	4.457821	Early > Late	Late > Early
L	Cingulate Gyrus	24	-2_+6_+28	23	4.309862	Early > Late	Late > Early
L	Cerebellar Crus I/VI	-	-9_-80_-18	16	3.781935	Early > Late	NS
L	Superior Frontal Gyrus	6	-1_+18_+64	33	4.311828	Early > Late	NS
L	Caudate/Putamen	-	-16_+12_+3	84	4.654485	Early > Late	NS
R	Caudate/Putamen	-	+20_+16_+1	49	4.429017	Early > Late	NS
R	Superior Parietal Lobule	7	+32_-55_+51	38	3.922614	Early > Late	NS
L	Middle Frontal Gyrus	10	-31_+53_+3	14	3.42056	NS	Early > Late
L	Cuneus	18	-8_-100_+3	18	3.984	NS	Early > Late
R	Lingual Gyrus	18	+20_-98_-5	36	4.525875	NS	Early > Late

In sum, the learning phase analyses revealed several activation patterns that differed between early and late learning. During early learning, a number of regions including bilateral caudate activated more strongly for high- than low-probability choices among both patients and controls, which was not apparent when collapsing across learning phase. Further, several cognitive control regions were activated more strongly in controls than in patients overall during early learning. When early and late learning are compared, dorsal striatum ROIs developed differential activation for A/C/E greater than B/D/F choices during late learning. In addition, regions including bilateral caudate activated more strongly during early than late learning in controls, with no difference in patients (primarily due to lower activation during the early phase).

Modeling Analysis: Q-values. In addition to providing information about learning rates as discussed in the behavioral analysis section, the Q-learning algorithm is also able to provide trial-by-trial estimates of expected value (Q values) and prediction error. Here, we wished to determine whether the expected value of choosing a given stimulus was related to brain activity at the time that the choice is made. As described in Methods, we created GLMs in which the Q value of the chosen stimulus was included as a trial-by-trial parametric modulator, and then conducted whole-brain ANOVAs by time to identify regions in which this modulation was significant. In this analysis, regions demonstrating a significant effect of time for the Q-value regressor are regions whose trial-by-trial variation in activation was captured by trial-by-trial variation in the expected value of the chosen stimulus. The modulatory effect may be conceptualized as a correlation between Q value and activation on each trial; for positive values, activation will be greater (or deactivation smaller) for larger Q values, and for negative values,

activation will be smaller (or deactivation greater) for larger Q values. As in the behavioral modeling analysis, this analysis was restricted to the "learners" group in order to avoid drawing conclusions from poorly fitting models.

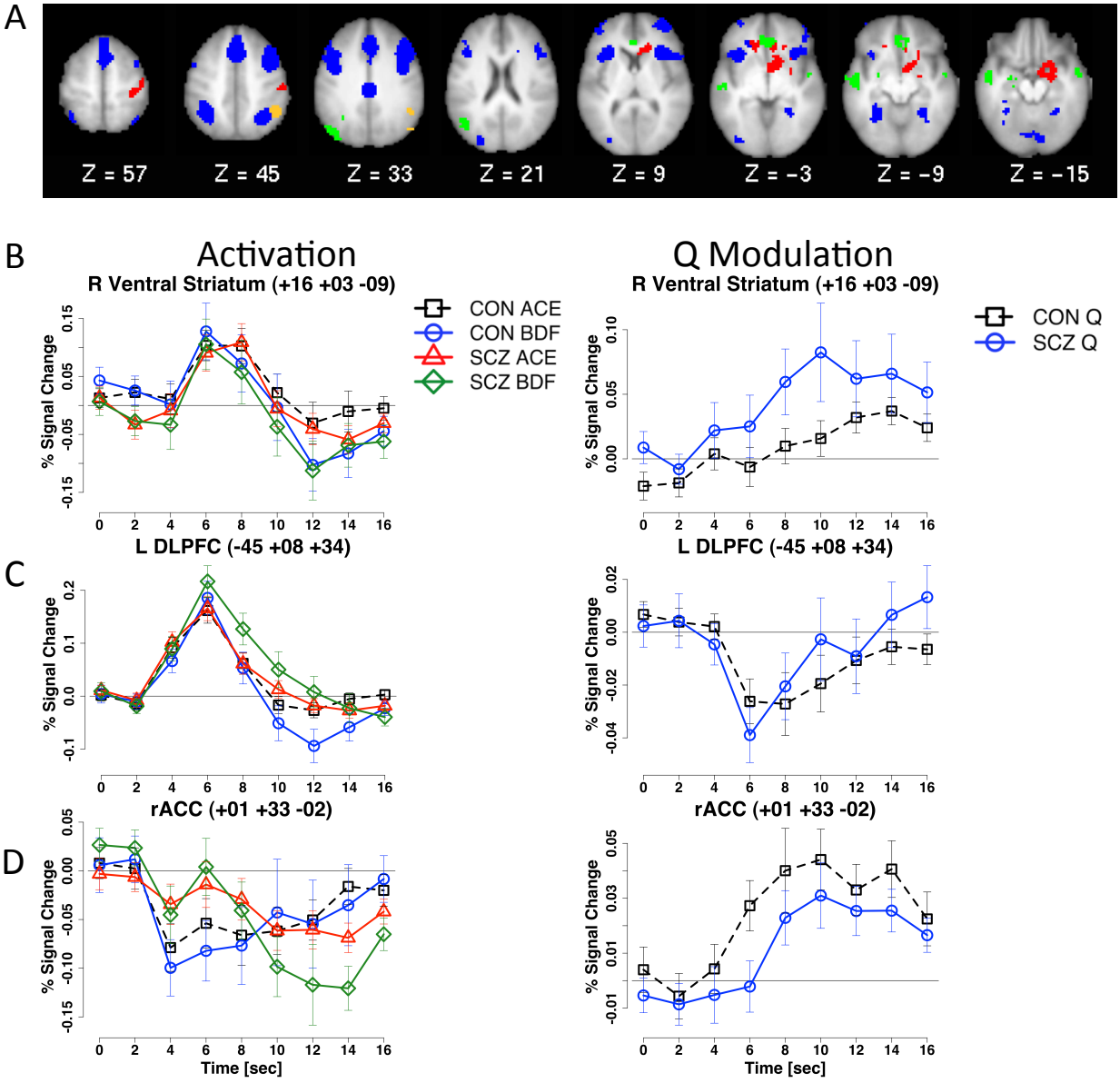
Striatal ROI analysis of Q value-related activity revealed significant positive modulation in bilateral nucleus accumbens (L: $F(8,416) = 6.92$, $p < 10^{-7}$) , R: $F(8,416) = 5.89$, $p < 10^{-8}$). This effect was significant within each group individually, and no group interactions were present. These regions demonstrated activation that was positively modulated, indicating greater activation on trials with higher Q values, in both patients and controls. These regions were also identified as part of the whole-brain analysis, for which example timecourses are shown in Figure 12b.

Whole-brain regions demonstrating a significant parametric effect of Q value among both groups combined are shown in Figure 12a and Table 7; for clarity, only regions that were significant on post-hoc tests within both patients and controls separately are discussed here. Several distinct subsets of regions showing similar activation patterns were identified. Similar to the ROI analysis results, right ventral striatum demonstrated a pattern of overall activation that was modulated positively by Q value, as did regions in right caudate, right postcentral gyrus, and left inferior frontal gyrus. As shown in Figure 12b, the magnitude of the Q value effect was larger in patients than in controls for the ventral striatum region, particularly at the time of peak activation; however, no significant interaction with group was present. A second pattern was seen in cognitive control network regions including bilateral DLPFC, posterior parietal cortex, anterior insula, and ACC/preSMA, among other regions. These regions had a pattern of overall activation with negative modulation, meaning that the lower the Q value of the chosen stimulus, the greater the activation in these regions. As shown in Figure 12c, the activation patterns and magnitudes were highly similar between patients and controls for all of these regions. Other activation patterns identified in this analysis included regions whose deactivation was reduced for higher Q values, including rACC (Figure 12d), and one region in inferior parietal lobule whose deactivation was increased for higher Q values.

Table 7. Regions demonstrating significant Q value modulation.

Region		BA	Talairach	# Voxels	Z	Activation Pattern	
						Choice vs Baseline	Q value modulation
R	Postcentral Gyrus	3	+42_-25_+54	82	4.79	Activation	Positive
R	Ventral Striatum	-	+16_+03_-09	220	6.98	Activation	Positive
R	Caudate	-	+11_+24_+03	116	5.69	Activation	Positive
L	Inferior Frontal Gyrus	47	-17_+29_-05	60	5.11	Activation	Positive
R	ACC/pSMA	8	+01_+19_+44	570	7.38	Activation	Negative
L	Anterior Insula	13	-32_+18_+03	251	6.87	Activation	Negative
R	Anterior Insula	44	+43_+18_+07	376	7.38	Activation	Negative
L	Fusiform Gyrus	37	-34_-56_-06	70	4.57	Activation	Negative
R	Cerebellar Crus I	-	+12_-85_-17	57	5.18	Activation	Negative
R	Cerebellar Vermis	-	+03_-72_-37	14	4.66	Activation	Negative
R	Lingual Gyrus	19	+29_-59_-03	65	4.87	Activation	Negative
R	Middle Frontal Gyrus	8	+42_+11_+41	598	6.28	Activation	Negative
L	Middle Frontal Gyrus	9	-45_+08_+34	232	5.03	Activation	Negative
R	Middle Frontal Gyrus	10	+39_+49_+07	155	7.41	Activation	Negative
L	Middle Frontal Gyrus	10	-42_+50_+05	74	6.56	Activation	Negative
L	Middle Frontal Gyrus	9	-47_+31_+27	92	4.52	Activation	Negative
L	Middle Occipital Gyrus	18	-30_-87_+06	85	4.54	Activation	Negative
L	Posterior Cingulate	23	-01_-26_+32	82	5.43	Activation	Negative
L	Posterior Parietal Cortex	40	-40_-53_+47	236	6.06	Activation	Negative
R	Posterior Parietal Cortex	7	+32_-61_+49	192	5.27	Activation	Negative
L	Cerebellar Crus II	-	-10_-87_-31	147	7.26	Activation	Negative
L	Superior Frontal Gyrus	6	-02_+10_+68	85	5.01	Activation	Negative
L	Middle Temporal Gyrus	21	-53_-17_-09	109	5.21	Deactivation	Positive
L	Middle Temporal Gyrus	39	-40_-70_+28	129	4.52	Deactivation	Positive
R	Middle Temporal Gyrus	21	+53_-08_-14	28	4.71	Deactivation	Positive
L	Parahippocampal Gyrus	28	-21_-13_-18	27	3.44	Deactivation	Positive
R	Anterior Cingulate	24	+01_+33_-02	129	6.60	Deactivation	Positive
R	Inferior Parietal Lobule	40	+49_-54_+35	127	4.80	Deactivation	Negative

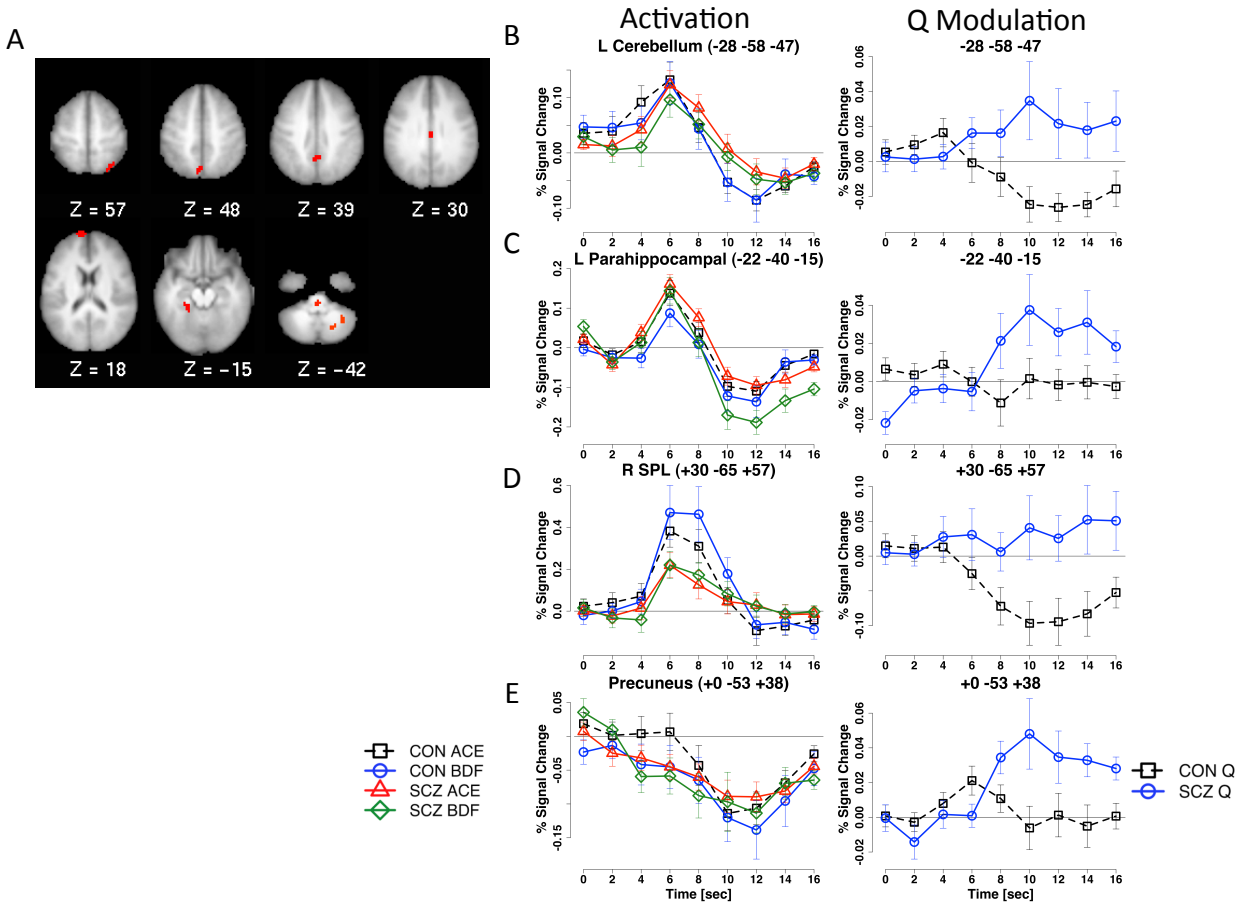
Figure 12: Q value effects. (A) Regions demonstrating significant Q-value modulation. Red = activation with positive modulation; blue = activation with negative modulation; green = deactivation with positive modulation; orange = deactivation with negative modulation. (B) Timecourse for right ventral striatum showing activation with positive modulation. (C) Timecourse for DLPFC showing activation with negative modulation. (D) Timecourse for rACC showing deactivation with positive modulation



Comparatively few regions had a significant Time X Group interaction in the Q value effect (Figure 13a). Regions in brainstem, bilateral cerebellum, and left superior frontal gyrus displayed activation that was negatively modulated in controls but positively modulated in patients (Figure 13b). Similarly, left parahippocampal gyrus was modulated positively in patients, but not in controls (Figure 13c), and right superior parietal lobule was modulated negatively in controls, but not in patients (Figure

13d). Finally, precuneus showed deactivation that was negatively modulated in patients, but not controls (Figure 13e).

Figure 13: Group interactions in Q value effect. (A) Regions with significant Q value x Time x Group effects. (B)-(D) Example timecourses as described in text.



Choice Analysis Summary. When the ANOVA and modeling analyses are taken together, a number of consistent patterns emerge. First, when activity from across the full acquisition phase is examined, activation patterns in patients and controls show a great deal of similarity, and comparatively few differences, particularly among cognitive control and reward-related regions. Second, across the full acquisition phase, differential choice-related activity is almost exclusively driven by regions activating more strongly for low-probability than high-probability choices. For example, the cognitive control network shows both strong activation for B/D/F choices and trial-by trial increases in activity for choices with lower

Q values, among both patients and controls. This pattern seems to be driven most strongly by the AB pair, and seems to be stronger late in learning.

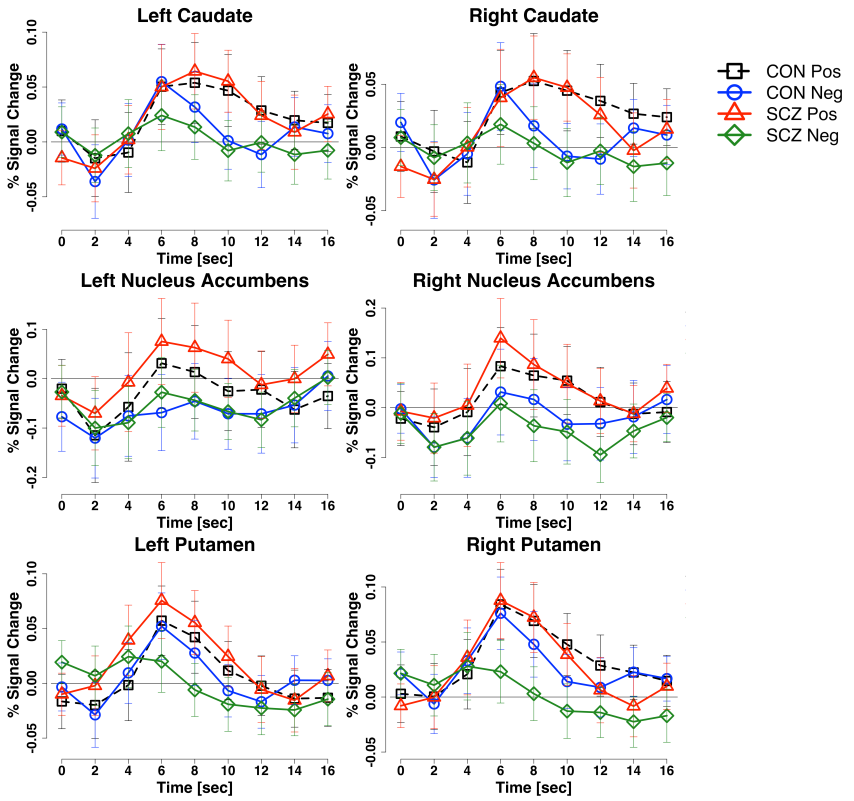
Regions activating more strongly for high- than low-probability choices emerge when analysis is restricted to early learning, and include striatal regions and orbitofrontal cortex. Striatal regions also demonstrate trial-by-trial increases in activation when choosing stimuli with higher Q values. There was no evidence of a reduction in these effects among patients; in fact, while there were no significant group differences, the main effects were often driven more strongly by the patient group.

Despite these similarities, a few group differences were apparent. Across the full acquisition phase, activation in cognitive and error-related regions was reduced in patients when making "B" choices, with lower activity associated with lower performance on the AB pair. Further, during early learning, patients activated the cognitive control network to a lesser extent than controls overall. When activation was compared across learning phases, group differences in striatal activity were present, with controls showing a reduction in activation across learning, while patients showed little activation in either phase.

Feedback-Related Activation

ANOVA analysis 1: Whole acquisition phase, collapsed across stimulus pair. The first ANOVA analysis is equivalent to that described for choice-related activity, and examined functional activation using a Choice (A/C/E, B/D/F) x Time (1-9) x Feedback (Positive, Negative) ANOVA for trials across the whole acquisition phase. In the striatal ROI analysis, all regions demonstrated significant Feedback x Time interactions that survived correction for multiple comparisons (all p values < .02). Example timecourses are shown in Figure 14, and demonstrate that both groups activated bilateral caudate, putamen, and nucleus accumbens more strongly for positive than negative feedback. There was no evidence of a reduction in this effect among patients as compared to controls. In fact, activation magnitudes for positive versus negative feedback in these regions were greater in patients than controls, although no Feedback x Time x Group interactions survived correction for multiple comparisons.

Figure 14. Feedback ANOVA ROI results showing Feedback x Time interactions in all regions.



A large number of regions outside the striatum also displayed a significant feedback by time interaction, as shown in Figure 15a. Only regions whose interaction was significant within each group separately are included here. A subset of these regions, including bilateral ventral striatum, caudate, and amygdala activated more strongly for positive than negative feedback in both patients and controls (Table 8a). A second subset of regions demonstrated greater activation for negative than positive feedback, and included ACC/preSMA, bilateral anterior insula, and bilateral superior parietal lobule (Table 8b). A similar activation pattern, in which peak activation was greater for negative than positive feedback, but activation for positive feedback was sustained longer than for negative feedback, was seen in regions including bilateral DLPFC, precentral gyrus, thalamus, and dorsal ACC (Table 8c). Finally, a set of regions including rACC and middle cingulate gyrus displayed deactivation that was greater for negative than positive feedback (Table 8d). Table 8 also includes post-hoc analyses within each stimulus pair, and shows that most of these relationships were driven by the CD and EF pairs more than the AB pair.

Table 8. Feedback ANOVA: Feedback x Time Interaction.
(a) Regions with greater activation for positive than negative feedback.

	Region	BA	Talairach	Voxels	Pattern	Z	AB	CD	EF
L	Amygdala	-	-23, -05, -11	77	Pos > Neg	5.38	NS	*	**
R	Caudate	-	+16, +12, +17	113	Pos > Neg	5.55	*	***	***
L	Caudate	-	-22, -09, +24	33	Pos > Neg	4.07	NS	***	*
R	Cuneus	19	+17, -88, +29	154	Pos > Neg	4.79	NS	*	**
R	Inferior Frontal Gyrus	46	+44, +37, +12	55	Pos > Neg	4.89	NS	NS	**
L	Inferior Parietal Lobule	40	-59, -34, +29	131	Pos > Neg	6.30	*	***	***
R	Inferior Parietal Lobule	40	+55, -27, +34	93	Pos > Neg	4.89	NS	**	**
L	Inferior Parietal Lobule	40	-51, -53, +50	16	Pos > Neg	4.86	NS	NS	*
L	Middle Occipital Gyrus	18	-29, -89, +05	92	Pos > Neg	4.96	NS	**	**
L	Middle Occipital Gyrus	18	-38, -82, -08	28	Pos > Neg	4.81	**	NS	**
L	Middle Temporal Gyrus	39	-45, -72, +20	46	Pos > Neg	4.56	NS	**	*
L	Middle Temporal Gyrus	21	-55, -15, -05	42	Pos > Neg	4.81	*	**	***
R	Middle Temporal Gyrus	21	+56, -08, -04	25	Pos > Neg	4.85	NS	NS	***
L	Middle Temporal Gyrus	37	-52, -53, +01	13	Pos > Neg	3.99	NS	**	**
L	Parahippocampal Gyrus	35	-21, -24, -11	99	Pos > Neg	5.86	NS	***	**
R	Parahippocampal Gyrus	30	+15, -32, -03	34	Pos > Neg	5.64	NS	***	***
L	Postcentral Gyrus	3	-59, -17, +33	55	Pos > Neg	4.98	*	**	***
L	Posterior Cingulate	29	-04, -41, +07	172	Pos > Neg	6.55	NS	***	***
R	Posterior Cingulate	30	+07, -52, +18	147	Pos > Neg	4.97	NS	***	***
R	Precentral Gyrus	6	+57, -10, +36	87	Pos > Neg	5.46	NS	**	**
R	Putamen	-	+28, -16, -08	129	Pos > Neg	5.72	NS	**	**
L	Superior Temporal Gyrus	41	-58, -26, +08	59	Pos > Neg	5.19	NS	NS	**
R	Superior Temporal Gyrus	42	+57, -36, +14	105	Pos > Neg	5.04	NS	NS	***
L	Superior Temporal Gyrus	41	-42, -37, +13	40	Pos > Neg	4.57	NS	**	**
R	Transverse Temporal Gyrus	41	+53, -17, +08	64	Pos > Neg	4.58	NS	NS	**
R	Ventral Striatum	-	+20, +04, -01	200	Pos > Neg	8.39	**	***	***
L	Ventral Striatum	-	-17, +07, +00	183	Pos > Neg	6.33	NS	***	***

Table 8b: Regions with greater activation for negative than positive feedback.

	Region	BA	Talairach	Voxels	Pattern	Z	AB	CD	EF
	ACC/preSMA	32	+01, +14, +45	359	Neg > Pos	5.15	NS	**	***
L	Cuneus	18	-11, -85, +12	226	Neg > Pos	5.12	NS	**	***
L	Inferior Frontal Gyrus	47	-46, +14, -04	20	Neg > Pos	4.72	NS	*	NS
L	Inferior Parietal Lobule	40	-38, -39, +41	32	Neg > Pos	4.23	NS	**	*
L	Anterior Insula	13	-34, +18, +07	97	Neg > Pos	4.86	NS	*	**
R	Anterior Insula	13	+32, +17, +11	49	Neg > Pos	4.18	NS	*	*
L	Lingual Gyrus	18	-12, -73, -05	108	Neg > Pos	4.84	NS	**	***
R	Lingual Gyrus	17	+06, -84, +03	126	Neg > Pos	4.46	NS	**	**
R	Superior Frontal Gyrus	6	+07, +07, +60	212	Neg > Pos	5.93	NS	***	***
L	Superior Parietal Lobule	7	-30, -52, +40	36	Neg > Pos	4.13	NS	NS	**
R	Superior Parietal Lobule	7	+28, -57, +44	65	Neg > Pos	4.11	NS	NS	***

Table 8c: Regions with greater activation for negative than positive feedback at peak, but sustained positive responses

	Region	BA	Talairach	Voxels	Pattern	Z	AB	CD	EF
L	Cingulate Gyrus	24	-07, -10, +41	93	Pos Sustained	4.19	NS	**	**
L	Cuneus	19	-14, -81, +30	277	Pos Sustained	5.05	NS	***	***
L	Cuneus	18	-03, -74, +21	208	Pos Sustained	4.91	NS	***	**
L	Fusiform Gyrus	37	-24, -64, -13	23	Pos Sustained	4.33	NS	**	NS
R	Insula	13	+46, +12, -03	44	Pos Sustained	4.75	NS	**	**
L	Insula/Central Operculum	13	-42, -15, +20	39	Pos Sustained	4.65	NS	*	*
R	Lingual Gyrus	18	+25, -74, -10	77	Pos Sustained	4.34	*	*	**
L	Lingual Gyrus	18	-25, -59, +07	62	Pos Sustained	4.82	NS	***	**
L	Medial Frontal Gyrus	6	-05, -18, +62	36	Pos Sustained	4.66	NS	*	***
L	Middle Frontal Gyrus	9	-44, +20, +33	59	Pos Sustained	4.12	NS	*	**
R	Middle Occipital Gyrus	19	+36, -67, +13	57	Pos Sustained	4.31	NS	*	***
R	Middle Temporal Gyrus	37	+50, -62, +10	29	Pos Sustained	4.24	NS	NS	**
L	Postcentral Gyrus	3	-40, -21, +36	152	Pos Sustained	5.37	NS	***	*
L	Posterior Cingulate	30	-09, -61, +14	249	Pos Sustained	5.92	NS	***	***
R	Precentral Gyrus	4	+44, -13, +47	222	Pos Sustained	5.66	NS	**	**
R	Precentral Gyrus	6	+58, -01, +11	70	Pos Sustained	5.88	NS	***	**
L	Precentral Gyrus	4	-46, -08, +47	209	Pos Sustained	5.48	NS	***	***
L	Precentral Gyrus	6	-31, -01, +33	97	Pos Sustained	4.57	NS	*	***
R	Sub-Gyral	40	+27, -40, +54	131	Pos Sustained	4.83	NS	NS	***
L	Superior Frontal Gyrus	6	-13, -03, +63	164	Pos Sustained	5.46	*	**	***
R	Superior Occipital Gyrus	19	+31, -78, +24	95	Pos Sustained	4.40	NS	NS	***
L	Superior Parietal Lobule	7	-22, -48, +60	155	Pos Sustained	5.45	NS	NS	***
R	Supramarginal Gyrus	40	+44, -38, +29	98	Pos Sustained	5.53	NS	*	***
R	Thalamus	-	+02, -26, +12	99	Pos Sustained	5.09	*	***	***

Table 8d: Regions with greater deactivation for negative than positive feedback

	Region	BA	Talairach	Voxels	Pattern	Z	AB	CD	EF
	rACC	32	-01, +42, +05	256	(-) Neg > Pos	6.47	**	***	***
	rACC	24	-02, +28, -03	98	(-) Neg > Pos	5.58	**	**	**
L	Cingulate Gyrus	31	-11, -27, +45	30	(-) Neg > Pos	4.15	NS	**	*
L	Angular Gyrus	19	-39, -72, +34	23	(-) Neg > Pos	4.07	NS	**	NS

Figure 15. Feedback ANOVA: Feedback x Time interactions. (A) Regions with significant Feedback x Time interactions. Red = Positive > Negative; Blue = Negative > Positive; Green = Deactivation, Neg > Pos; Purple = Negative > Positive at peak, with sustained positive response. (B) Example timecourses for each response pattern.

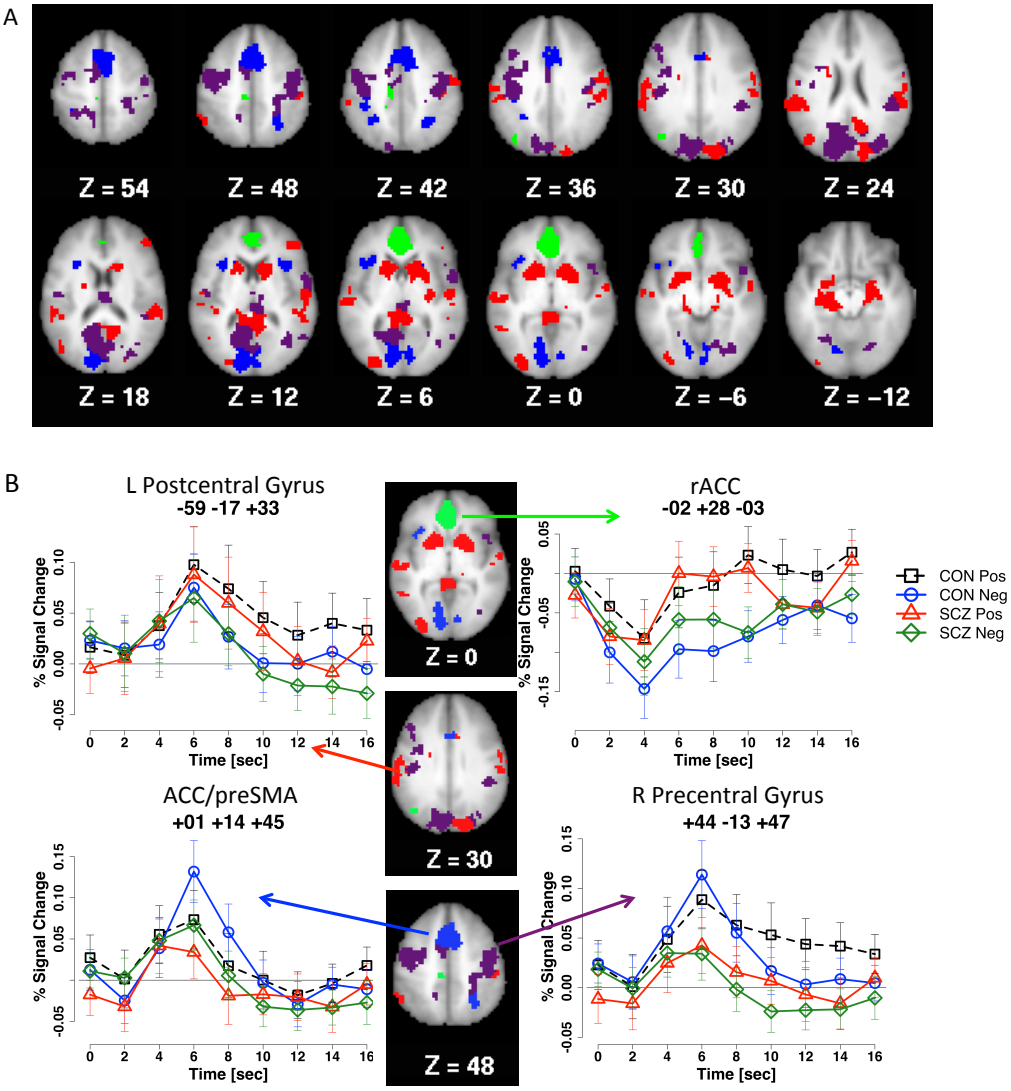


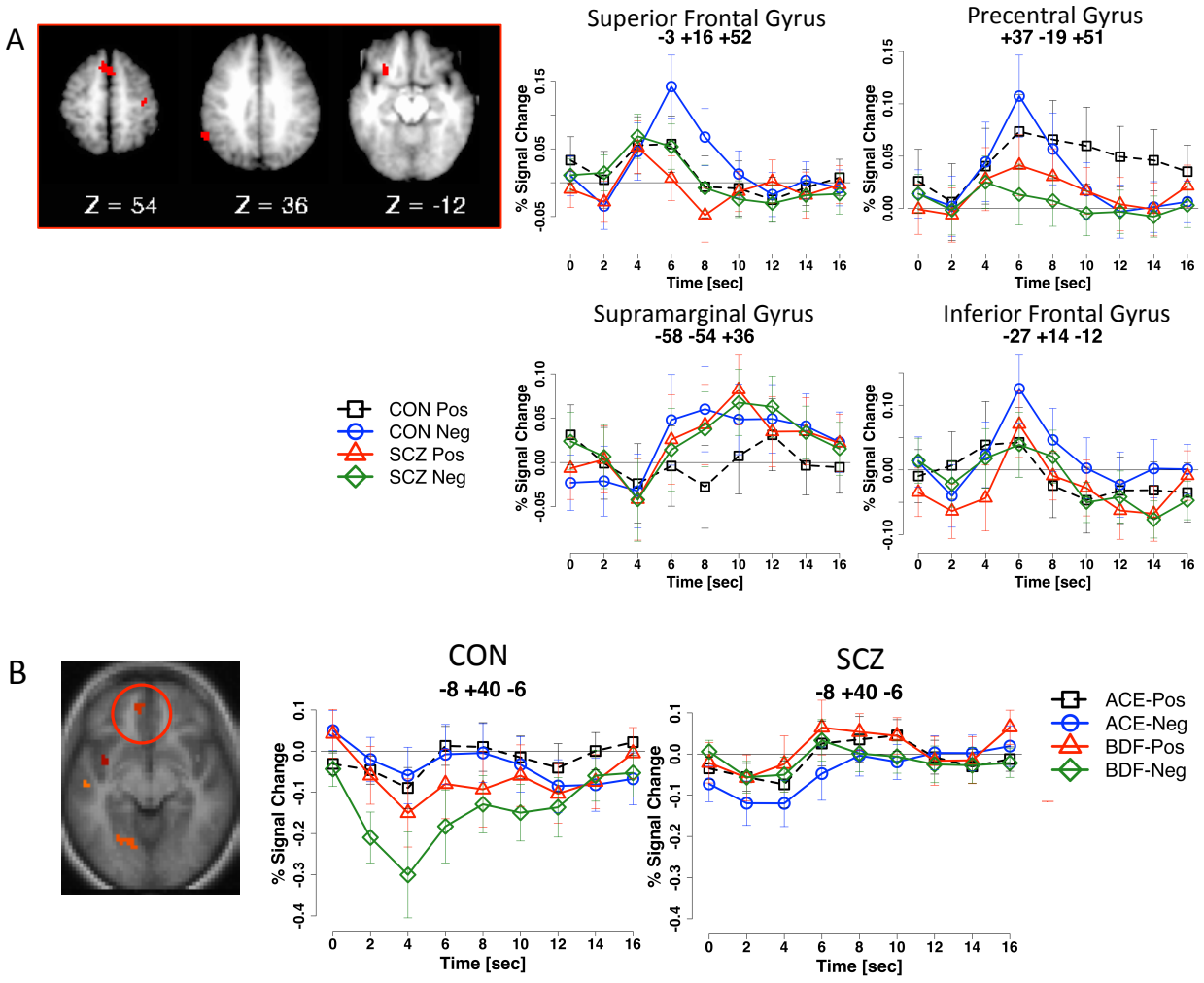
Table S2 and Figure S1 in the supplement describe additional effects and interactions found in the feedback ANOVA analysis that did not interact with group. A set of regions including several members of the cognitive control network (left DLPFC, PPC, anterior insula, ACC/preSMA, and right dorsal premotor cortex) activated more strongly to feedback received after high-probability choices, regardless of feedback type, yielding significant Choice x Time interactions. In addition, one region in right middle frontal gyrus showed a Choice x Feedback x Time interaction with reduced activation for low-probability choices that received positive feedback as compared to all other conditions.

Table 9 describes effects and interactions found in the feedback ANOVA analysis that did interact with group. A Time x Group interaction with greater activity in controls than patients irrespective of feedback type was seen in regions including right inferior parietal lobule and left middle frontal gyrus, while greater activity for patients than controls was seen in occipital regions. Figure 16a shows the few regions that displayed a Feedback x Time x Group interaction. These include regions in left superior and inferior frontal gyri and right precentral gyrus that activated more strongly for negative than positive feedback in controls, with less differentiation between conditions in patients. There were also a few regions displaying a significant Choice x Time x Group interaction, whose activation patterns differed between groups for high- versus low-probability choices regardless of feedback. These regions included R superior frontal gyrus, which activated for BDF choices in controls and ACE choices in patients; cerebellar crus I which showed the opposite patterns, and right angular gyrus and inferior frontal gyrus, which activated for A/C/E choices in patients, but not in controls. Finally, there was a Choice x Feedback x Time x Group interaction in a small set of regions including left insula and VMPFC. Activation timecourses for the VMPFC region are shown in Figure 16b, and reveal deactivation that was strongest for low-probability choices given negative feedback among controls, with little differentiation among conditions in patients.

Table 9: Feedback ANOVA: Interactions with Group

Effect		Region	BA	Talairach	Voxels	Z	Activation Pattern	
							CON	SCZ
Time x Group	R	Lingual Gyrus	19	+18_-61_-6	40	4.31	CON > SCZ	
	L	Cuneus	19	-16_-87_+26	106	5.12	CON > SCZ	
	R	Inferior Parietal Lobule	40	+62_-38_+24	18	4.77	CON > SCZ	
	L	Lingual Gyrus	19	-15_-59_-1	69	4.62	CON > SCZ	
	L	Middle Frontal Gyrus	10	-27_+51_+21	93	5.27	CON > SCZ	
	L	Middle Temporal Gyrus	22	-62_-41_+5	25	4.64	CON > SCZ	
	L	Fusiform Gyrus	19	-25_-89_-18	32	5.15	SCZ > CON	
	R	Fusiform Gyrus	19	+25_-79_-17	19	4.65	SCZ > CON	
	R	Fusiform Gyrus	37	+37_-46_-15	26	4.40	SCZ > CON	
	L	Middle Occipital Gyrus	37	-43_-71_+1	16	4.36	SCZ > CON	
	L	Parahippocampal Gyrus	28	-18_-16_-17	24	4.58	SCZ > CON	
	R	Cuneus	19	+18_-88_+24	144	5.40	SCZ > CON at peak; CON sustained	
	R	Insula	21	+45_-3_-8	21	4.67	SCZ > CON at peak; CON sustained	
	L	Superior Temporal Gyrus	22	-46_-22_+0	25	4.24	SCZ > CON at peak; CON sustained	
	R	Angular Gyrus	39	+57_-63_+34	14	4.39	(-) SCZ > CON	
FB x Time x Group	L	Inferior Frontal Gyrus	13	-27_+14_-12	37	5.04	Neg > Pos	Pos > Neg
	L	Superior Frontal Gyrus	8	-3_+16_+52	42	4.07	Neg > Pos	Neg > Pos
	L	Supramarginal Gyrus	40	-58_-54_+36	17	4.49	Neg > Pos	NS
	R	Precentral Gyrus	4	+37_-19_+51	16	3.83	Neg > Pos at peak; Pos sustained	NS
Choice x Time x Group	L	Genu of Corpus Callosum	-	-15_+25_-3	19	4.46	(-) BDF > ACE	(-) ACE > BDF
	R	Medial Temporal White Matter	-	+38_-33_-5	134	6.08	ACE > BDF	BDF > ACE
	R	Cerebellar Crus I	-	+52_-53_-28	30	5.66	ACE > BDF	BDF > ACE
	R	Superior Frontal Gyrus	10	+29_+60_+4	43	4.93	BDF > ACE	ACE > BDF
	R	Angular Gyrus	40	+56_-57_+36	32	5.28	NS	(+) ACE, (-) BDF
	R	Inferior Frontal Gyrus	46	+52_+37_+11	17	5.37	NS	(+) ACE, (-) BDF
Choice X FB x Time x Group	L	Insula	13	-40_+3_-13	24	4.21	BDF-Neg > ACE > BDF-Pos	BDF-Neg > BDF-Pos > ACE
	L	Lingual Gyrus	18	-19_-66_-10	17	3.69	NS	BDF-Neg > ACE > BDF-Pos
	L	Superior Temporal Gyrus	22	-50_-20_-6	22	4.21	NS	BDF-Pos > BDF-Neg > ACE
	L	Ventromedial PFC	10	-8_+40_-6	16	4.56	(-) BDF-Neg > BDF-Pos > ACE	NS

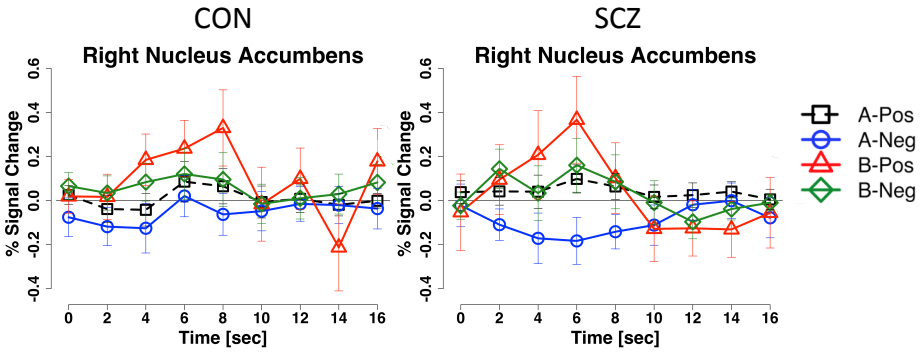
Figure 16. Feedback ANOVA: Interactions with Group. (A) Regions with Feedback x Time x Group interactions. (B) VMPFC region with a Choice x Feedback x Time x Group interaction.



Overall, the feedback ANOVA revealed robust main effects of feedback that did not differ between groups in most regions associated with reward or cognitive control. Striatal regions activated more strongly for positive than negative feedback, while cognitive control regions showed the opposite pattern. Several cognitive control regions also demonstrated greater activation for high- than low-probability choices at the time of feedback, regardless of the feedback actually received, in both groups. In terms of group effects, a few cortical regions demonstrated reduced feedback responses overall, or reduced responses to negative vs. positive feedback, while a set of cerebellar regions activated more strongly in patients overall, and VMPFC showed responses that varied with both choice and feedback among controls, but were absent in patients. However, no evidence was found for altered striatal responses to feedback among patients.

Analysis 2: AB only. As in the analysis for choice-related activity, the AB pair was also examined alone using a repeated measures ANOVA with Choice (A, B), Feedback (Positive, Negative), and Time (1-9) as within-subjects factors and Group (CON, SCZ) as a between subjects factor. Striatal ROI analysis revealed significant feedback x time ($F(8,504) = 2.94$, $p < .004$) and choice x time ($F(8,504) = 3.59$, $p < .0005$) interactions in right ventral striatum, which were driven by elevated responses to "B" choices that yielded positive feedback (in essence, a positive prediction error). This activity was seen in both groups, and no significant interactions with group were present (Figure 17). The whole brain Feedback x Time activation map for the AB pair was much more sparse than that seen for all three pairs combined, with only the rACC and bilateral ventral striatal regions reaching significance. There were no additional regions demonstrating interactions with group that were not present in the analysis for all three pairs combined.

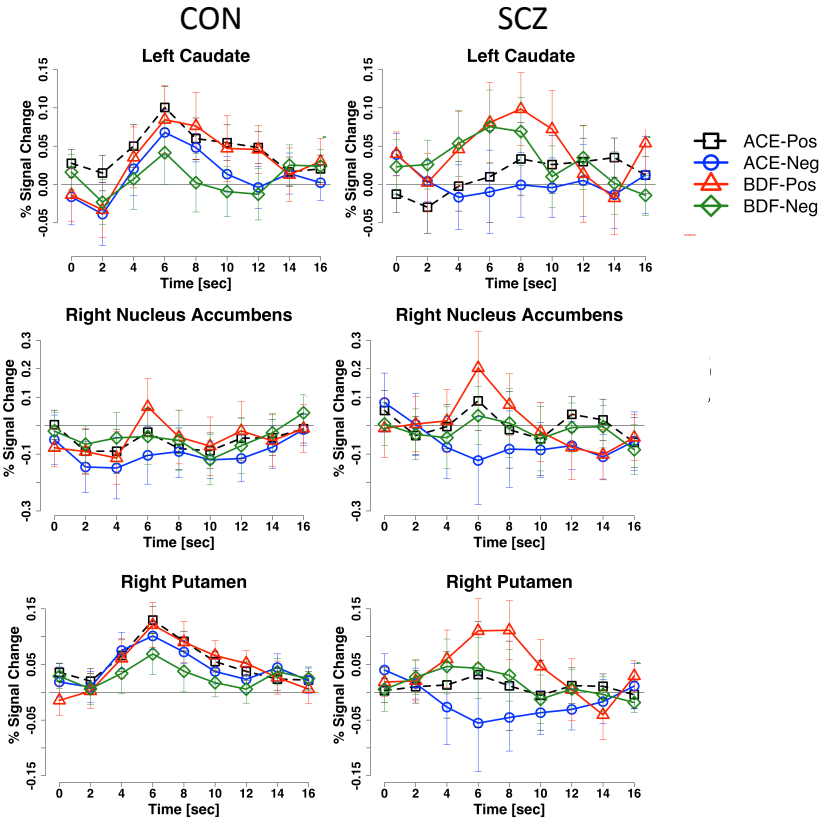
Figure 17. Feedback ANOVA, AB pair: ROI analysis. Example timecourses shown for nucleus accumbens ROI with significant Feedback x Time and Choice x Time interactions.



Analysis 3: Learning Phase. To examine whether responses to positive and negative feedback differed during the early stages of learning, the Choice x Feedback x Time x Group ANOVA was repeated within the early learning phase. ROI analysis revealed significant Feedback x Time interactions in right caudate, right nucleus accumbens, and bilateral putamen (all p values < .002). There was also a Choice x Feedback x Time x Group interaction in left caudate ($F(8,440) = 2.77$, $p<.006$), in which patients had stronger responses, especially to positive feedback, after low-probability choices, while controls responded to positive feedback after both choice types (Figure 18). While there were no interactions with choice in the other regions, the pattern of activity in right nucleus accumbens and bilateral putamen

resembled what one might expect for a prediction error signal, particularly in the patient group, with the strongest activation for unexpected positive outcomes (B/D/F choices with positive feedback; red) and deactivation for unexpected negative outcomes (A/C/E choices with negative feedback; blue). On whole-brain analysis, feedback effects during early learning closely mirrored those seen in the full acquisition phase; the major activation patterns were the same as those reported above, and no new meaningful effects emerged upon restriction of the analysis to early learning.

Figure 18. Feedback ANOVA within early learning phase, ROI analysis.



When learning phase was included as a factor, there were no significant effects or interactions with this factor in the striatal ROI analysis. The whole-brain ANOVA revealed a number of regions whose activation differed between early and late learning, however, which are described in Table S3. The only factor that interacted significantly with learning phase was choice, with several regions demonstrating a Learning Phase x Choice x Time interaction. The majority of these regions showed the development of an effect of choice on feedback-related activation during late learning that had not been present during

early learning (Figure S2). There were no results in the learning phase analysis that interacted with group.

Modeling Analysis: Prediction Errors. In addition to examining functional activation changes associated with positive and negative feedback, we wished to examine whether trial-by-trial fluctuations in activation at the time of feedback were correlated with prediction error values derived from the reinforcement learning model. As described previously, GLMs for this analysis included signed prediction error (i.e. positive and negative) as a parametric regressor, which was analyzed for significance using ANOVAs with time as a factor. A separate analysis also examined positive and negative prediction errors coded separately. As in the other modeling analyses, only "learners" were included to ensure that the prediction error terms were derived from well-fitting models. On striatal ROI analysis, all regions demonstrated significant prediction error effects (all p values < .007) with positive modulation, none of which interacted with group. These regions were also identified by the whole-brain analysis and are discussed further below.

Whole-brain regions with a significant prediction error effects are shown in Table 10 and Figure 19, and show a high degree of overlap with those regions with a feedback by time interactions in the ANOVA analyses. A set of regions demonstrating activation that was positively modulated (Table 10a), such that activation increased on trials with larger positive prediction errors, included bilateral ventral striatum and amygdala. Regions including rostral ACC and medial frontal gyrus demonstrated deactivation that positively modulated, and therefore reduced for larger positive prediction errors. A second subset of regions demonstrated activation with negative modulation, such that activation was higher for smaller positive (or larger negative) prediction errors (Table 10b). These regions included cognitive control regions such as bilateral DLPFC, PPC, anterior insula, and preSMA. Examples of each type of activation pattern are shown in Figure 19b.

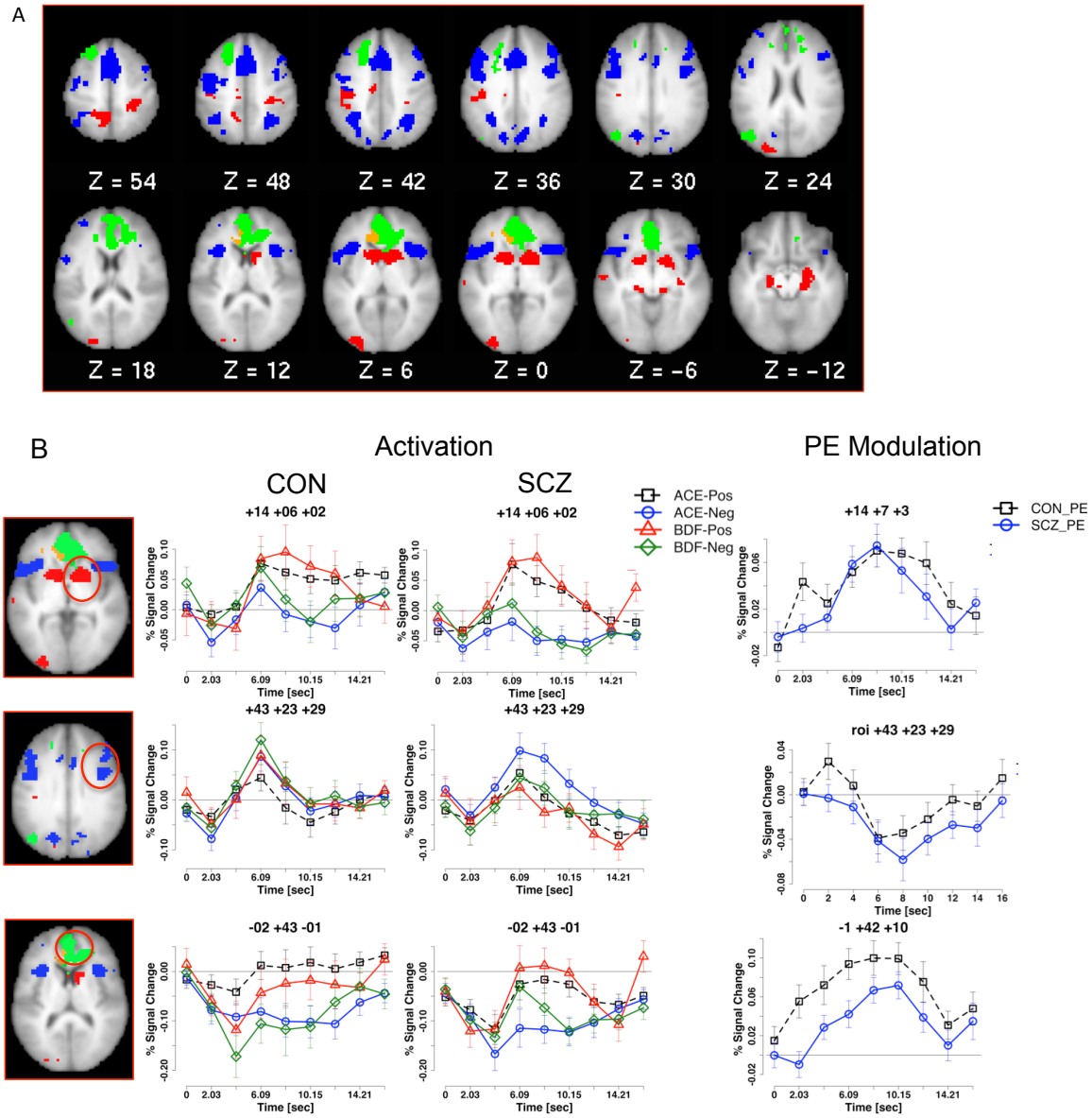
Table 10a. Prediction error analysis: regions demonstrating positive modulation in both groups.

	Region	BA	Talairach	Voxels	Z	Activation	PE modulation
L	Ventral Striatum	-	-11_+09_+02	167	5.94	Activation	Positive
L	Cingulate Gyrus	24	-11_-19_+40	27	4.16	Activation	Positive
R	Ventral Striatum	-	+14_+06_+02	198	6.71	Activation	Positive
R	Hippocampus	-	+29_-15_-09	60	4.87	Activation	Positive
L	Inferior Parietal Lobule	40	-42_-32_+41	102	4.53	Activation	Positive
L	Middle Occipital Gyrus	18	-26_-95_+04	77	5.17	Activation	Positive
L	Middle Occipital Gyrus	18	-22_-88_+22	45	4.10	Activation	Positive
L	Middle Temporal Gyrus	21	-61_-19_-03	15	4.23	Activation	Positive
R	Amygdala	35	+18_-28_-09	46	5.57	Activation	Positive
L	Amygdala	35	-15_-27_-09	57	5.19	Activation	Positive
R	Postcentral Gyrus	3	+29_-34_+55	107	4.58	Activation	Positive
R	Postcentral Gyrus	3	+18_-37_+69	43	4.13	Activation	Positive
L	Precuneus	7	-13_-48_+54	100	4.87	Activation	Positive
L	Anterior Cingulate	32	-02_+43_-01	254	8.12	Deactivation	Positive
R	Anterior Cingulate	32	+18_+34_+14	224	4.62	Deactivation	Positive
R	Anterior Cingulate	25	+06_+19_-05	120	5.58	Deactivation	Positive
R	Anterior Cingulate	24	+02_+25_+14	147	4.63	Deactivation	Positive
L	Anterior Cingulate	32	-19_+33_+23	32	3.90	Deactivation	Positive
L	Medial Frontal Gyrus	9	-03_+53_+15	177	6.38	Deactivation	Positive
L	Middle Frontal Gyrus	8	-25_+12_+39	53	4.56	Deactivation	Positive
L	Middle Temporal Gyrus	39	-44_-72_+25	77	4.29	Deactivation	Positive
L	Superior Frontal Gyrus	8	-21_+28_+46	148	6.00	Deactivation	Positive
L	Superior Frontal Gyrus	8	-33_+23_+54	28	4.26	Deactivation	Positive

Table 10b. Prediction error analysis: regions demonstrating negative modulation in both groups.

	Region	BA	Talairach	Voxels	Z	Activation	PE modulation
R	Cingulate Gyrus	32	+05_+18_+41	214	5.67	Activation	Negative
L	Cingulate Gyrus	32	-10_+19_+33	82	4.79	Activation	Negative
L	Inferior Frontal Gyrus	9	-47_+07_+32	201	4.50	Activation	Negative
L	Anterior Insula	13	-34_+17_+06	171	5.69	Activation	Negative
R	Anterior Insula	13	+40_+16_+06	185	6.23	Activation	Negative
R	Middle Frontal Gyrus	9	+43_+23_+29	143	5.18	Activation	Negative
R	Middle Frontal Gyrus	6	+37_+08_+53	72	4.43	Activation	Negative
L	Precentral Gyrus	4	-40_-11_+52	192	5.10	Activation	Negative
R	Precentral Gyrus	6	+40_+02_+35	130	4.77	Activation	Negative
L	Precentral Gyrus	9	-42_+25_+37	95	4.68	Activation	Negative
R	Precuneus	7	+07_-72_+38	94	4.02	Activation	Negative
L	Precuneus	7	-20_-71_+37	107	4.53	Activation	Negative
L	Superior Frontal Gyrus	6	-03_+10_+51	250	7.22	Activation	Negative
R	Superior Frontal Gyrus	6	+05_-01_+64	157	6.25	Activation	Negative
L	Superior Frontal Gyrus	6	-14_-01_+66	78	5.23	Activation	Negative
R	Superior Frontal Gyrus	6	+12_+14_+62	77	4.83	Activation	Negative
L	Superior Frontal Gyrus	10	-34_+49_+21	32	4.37	Activation	Negative
R	Superior Parietal Lobule	7	+31_-55_+43	142	5.21	Activation	Negative
L	Superior Parietal Lobule	40	-35_-52_+48	159	4.50	Activation	Negative
L	Superior Temporal Gyrus	22	-49_+10_+02	127	6.05	Activation	Negative
R	Superior Temporal Gyrus	38	+46_+12_-08	79	6.38	Activation	Negative
L	Anterior Cingulate	32	-14_+29_+08	97	4.79	Deactivation	Negative

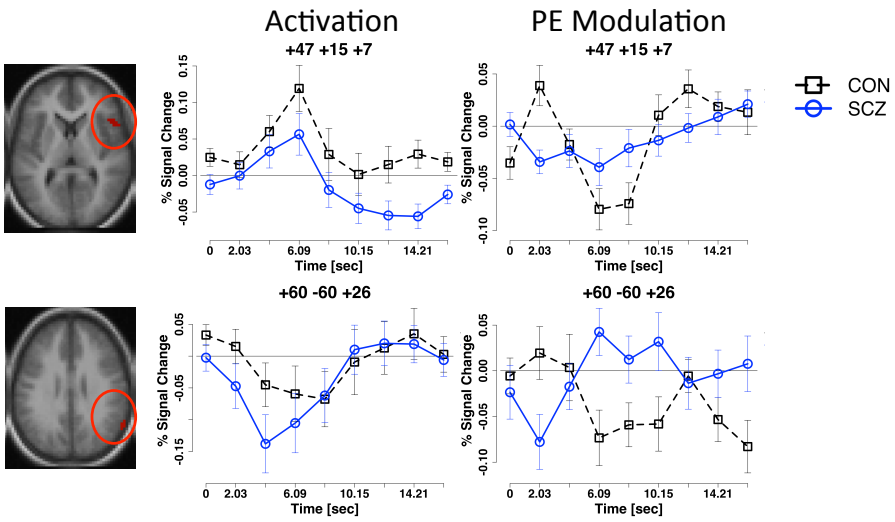
Figure 19. Prediction error analysis: Prediction Error x Time interactions. (A) Regions with significant prediction error effects in both groups. Red = activation with positive modulation; blue = activation with negative modulation; green = deactivation with positive modulation; orange = deactivation with negative modulation. (B) Example timecourses for each activation pattern.



In addition to the main effects of prediction error, there were two regions that demonstrated an interaction between prediction error and group (Figure 20). These regions included right inferior frontal gyrus and right superior temporal gyrus. Superior temporal gyrus showed deactivation that was negatively modulated in controls and positively modulated in patients, such that activation was increased for smaller positive or greater negative prediction errors in controls, but greater positive prediction errors in patients. Right inferior frontal gyrus showed activation that was negatively modulated in controls but

not patients, such that smaller positive or greater negative prediction errors yielded greater activation in controls only.

Figure 20: Prediction Error x Time x Group interactions



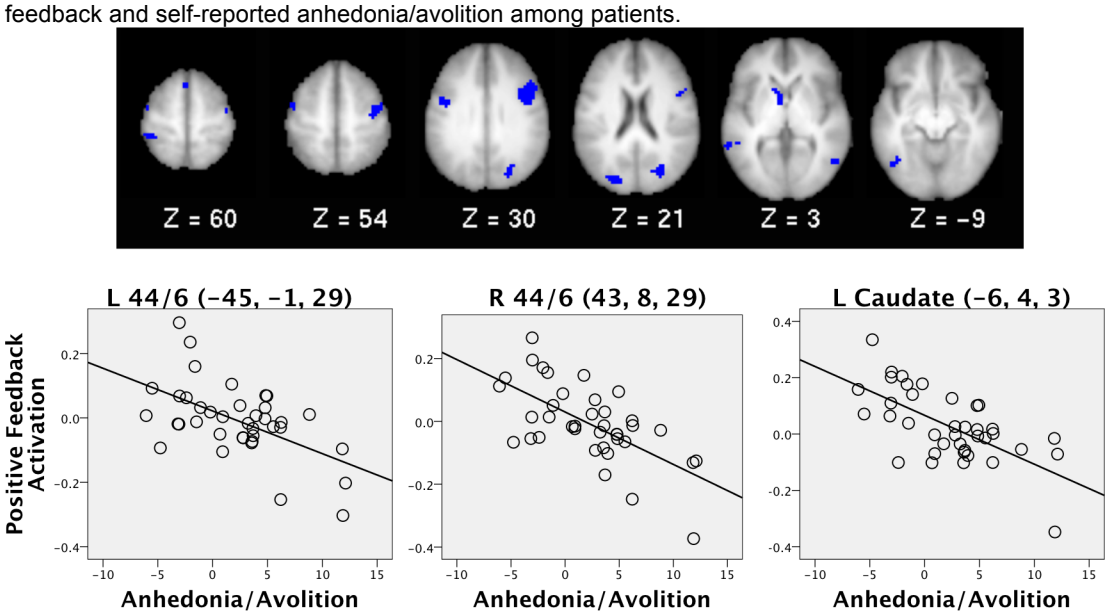
In order to determine whether positive and negative prediction errors were differentially affected in the patient group, we also examined positive and negative prediction errors separately. The positive prediction error effect was significant in bilateral nucleus accumbens (Left: $F(8,408)=3.86$, $p<.0003$; Right: $F(8,408)=2.71$, $p<.003$), while the negative prediction error effect was not significant in the striatal ROIs. No significant effects of group emerged on ROI or whole-brain analysis of positive and negative prediction errors separately.

Overall, activation at the time of feedback was highly similar between patients and controls. Both groups demonstrated greater activation of striatal regions for positive than negative feedback, and of cognitive control regions for negative more than positive feedback. Analysis of prediction error modulation revealed the same pattern. In addition, in both groups there was activity driven by the choice that had been made irrespective of the feedback received; this activity developed later in learning and was present in both groups. The only group difference identified in striatal regions was in left caudate, which demonstrated an unexpected pattern in which controls responded primarily to positive versus negative feedback, while patients showed a modulation of feedback response by choice type. Outside the striatum, VMPFC showed activation that was modulated by both choice and feedback in controls, but not

patients, though the pattern of activity was not consistent with prediction error coding. Overall, there were few robust group differences in regions associated with reward processing or learning.

Individual Differences Analysis. To determine whether there was a relationship between choice- or feedback-related activation and anhedonia/avolition severity, we conducted correlation analyses between activation averaged at timepoints 4 and 5 (peak activation) and the summary anhedonia/avolition scores derived from our clinical and self-report measures. First, we examined whether activity during high-probability choices was reduced in patients who were higher in anhedonia/avolition, using both striatal ROI and whole-brain analyses. This analysis was conducted within the early learning phase, because we had identified greater activity for high- than low-probability choices only during early learning. No significant results were obtained. Next, we examined whether anhedonia/avolition scores were associated with responses to positive feedback or positive prediction errors, using both striatal ROI and whole-brain analyses. There were no significant associations with positive prediction errors, and ROI analysis of responses to positive feedback did not survive correction for multiple comparisons. However, whole-brain analysis revealed significant negative relationships between anhedonia/avolition self-report scores and responses to positive feedback in several regions including left caudate and bilateral posterior DLPFC (Brodmann areas 44/6) (Figure 21). These regions showed reduced activation in response to positive feedback in those patients who were higher in anhedonia/avolition. When these same regions were examined in controls, no significant relationships were found. Lastly, we examined the regions that had demonstrated group differences in the above analyses to determine whether the effect that differed between groups further varied with anhedonia/avolition severity. This analysis failed to identify any significant relationships.

Figure 21. Correlation analyses: Regions with significant negative correlations between responses to positive feedback and self-reported anhedonia/avolition among patients.



Medication Effects. Given that the vast majority of our patients were taking antipsychotic medications that block dopamine receptors, we wished to examine whether there was a relationship between medication dose and striatal activation during either choice or feedback. To do this, we converted antipsychotic doses to standard chlorpromazine dose-equivalents (Gardner et al 2010), and conducted correlations between these dosages and our behavioral and neuroimaging measures of interest. Behaviorally, there was a significant relationship between medication dose and AvoidD (rho = .408, $p < .02$), wherein AvoidD scores were higher for larger medication doses. Dose did not correlate with ChooseC, gain or loss learning rate, or anhedonia/avolition. In the neuroimaging data, we examined all regions with significant group differences in the analyses above to determine whether the effect that differed between groups correlated with medication dose. None of the examined relationships were significant. Next, we conducted the same ROI and whole-brain correlations as described for the anhedonia/avolition analysis above, which also failed to yield significant results.

Discussion:

In this study, we asked a number of questions about whether probabilistic reinforcement learning was impaired in a sample of medicated individuals with schizophrenia. First, we asked whether patients had a specific deficit in learning from positive feedback, and whether this deficit was related to symptoms of anhedonia and avolition. Behavioral results using transfer measures sensitive to Go vs NoGo learning and gain versus loss learning rates from RL modeling suggested impairments in learning from positive, but not negative, feedback in patients as compared to controls. While the effect for the learning rates was less robust than that seen for the transfer measures, it is important to keep in mind that the sample for modeling analysis included only those subjects who demonstrated above-chance performance on the task, reducing the power and limiting the dynamic range of this analysis. It is also notable that the ChooseC and gain learning rate measures were correlated among both groups, supporting the claim that these measures both assess learning from positive feedback; however, the lack of the analogous correlation between AvoidD and loss learning rate, which has been demonstrated in past studies (Frank et al 2007), calls into question the extent to which these measures tapped into the same construct of NoGo learning in this sample.

While the behavioral results demonstrated the expected impairment in Go learning, we found no evidence in the neuroimaging results for a reduction in striatal responses among patients as compared to controls. We had hypothesized that Go learning impairments would be associated with reduced positive prediction error activity, and with reduced anticipatory reward responses (responses to high-probability choices), in striatal regions among patients. Instead, we found that striatal activity was intact in patients at the time of choice and feedback. We found no group differences in striatal activity for positive versus negative choices or feedback, expected values, or prediction errors when collapsing across the full acquisition phase. Separate analysis of positive and negative prediction errors also demonstrated intact striatal activation. During early learning, choice-related activity did not differ between groups, while at feedback, left caudate activated more strongly for positive feedback after any choice in controls, but for positive (and, to a lesser extent, negative) feedback after BDF choices in patients. This pattern is unexpected and contrary to our hypotheses, as it demonstrates a greater modulatory effect of choice on

activation at the time of feedback in patients than in controls. In fact, this and other striatal regions demonstrated activity that was if anything more characteristic of a prediction error signal in patients than in controls, although the group differences were not significant. The only suggestion of reduced striatal activation in patients as compared to controls was in the learning phase comparison, where bilateral caudate showed a greater reduction in overall activation at the time of choice between early and late learning in controls than in patients. This relationship was driven by lesser activation during early learning in patients as compared to controls, suggesting that patients may have a deficit in early recruitment of these regions. However, this effect did not differ by choice type, as would be expected if it represented a difference in reward anticipation; nor did it correlate with learning rate, overall performance, or clinical symptoms. Therefore, any claims about the importance of this activation difference in interpreting group differences in behavior or symptomatology are necessarily weak.

Our findings of largely intact striatal activation during learning contrast with the findings of several studies in the literature demonstrating reduced striatal prediction errors in medicated patients (Gradin et al 2011; Koch et al 2010; Morris et al 2012; Murray et al 2008; Waltz et al 2009). However, Walter et al (Walter et al 2009) also found intact striatal prediction errors in medicated schizophrenia patients, and the studies indicating reduced prediction errors have found these effects in mostly non-overlapping striatal subregions. Further, sample sizes in these studies were small (all Ns < 21, with two studies with N < 15), raising concern about the reliability of the reported effects. Another possible source of these differences are clinical differences in the populations examined; our patients had lower positive symptom severity than many published reports, and there is evidence in the literature that aberrant prediction error activity is related to positive symptoms in schizophrenia (Kapur 2003).

The lack of significant relationships between anhedonia/avolition scores and striatal prediction error activity reported here also differs from some reports in the literature (Morris et al 2012; Waltz et al 2009), although at least one other study looked for these relationships and did not find them (Gradin et al 2011). While our null findings may reflect a failure to detect an effect that exists in the population, it is worth noting that the present study is the only one in the literature with appropriate power to detect individual differences within the patient group. However, behavioral studies with larger samples have also reported relationships between reinforcement learning impairments and negative symptom severity

(Gold et al 2012; Waltz et al 2011), which we failed to detect in our behavioral results. We can only speculate that these differences may have been influenced by the specifics of the experimental design, as discussed below.

One important difference between our task and many in the literature is that patients received additional practice on the task before the scanning session. This was done to avoid an influence of confusion about task procedures, which was common among patients but not controls, on our results. We acknowledge that this procedure introduced a practice mismatch between groups, but wish to point out that mismatches in the amount of practice given are not unusual in prediction error studies, which sometimes have subjects train to criterion before entering the scanner. We gave a fixed number of practice trials because the goal of our practice session was not to establish conditioned associations, but to ensure familiarity with experimental procedures. Controls underwent a shorter practice session immediately prior to scanning, and were also administered additional practice trials in the event that they demonstrated confusion about task instructions. It is possible that this additional practice in patients contributed to the relative lack of group differences in our study as compared to others in the literature, which has interesting implications both clinically and for future studies in this population.

Given that explicit learning systems associated with prefrontal cortical function are also expected to contribute to learning in this task, we also examined measures of explicit learning and whole-brain activity. While the basal ganglia learning system supports the gradual integration of feedback over several trials, regions such as orbitofrontal cortex, prefrontal cortex, and medial temporal lobe are commonly associated with more rapid, explicit forms of learning. Orbitofrontal cortex is thought to be involved in a form of working memory for value, and is commonly implicated in tasks where reward contingencies must be updated more rapidly than can be accommodated by the basal ganglia (Frank and Claus 2006). Medial temporal lobe is associated with declarative memory, and may be involved in the earliest stages of implicit learning tasks that later rely on the basal ganglia, as well as in making inferences from learned associations (Poldrack et al 2001). Prefrontal cortex is thought to be involved in representing contextual information about the states of the environment and of the actor in order to exert top-down bias on choice behavior (Doya 1999). While this task was designed to rely heavily on the basal ganglia system, it is likely that these explicit systems contributed to learning as well, and may have

differed between groups. Our behavioral data provides some evidence that this may have been the case. At both acquisition and test, we found evidence of reduced performance on the AB pair in patients as compared to controls, while performance on the CD pair and EF pair (at test) did not differ between groups. We speculate that this finding may be related to an impairment in explicit learning among patients, given that the AB pair had the highest probability ratio and was therefore the easiest to learn via explicit mechanisms. This is because higher ratios require fewer trials to be held in working memory in order for explicit representations of reward contingencies to be formed, while lower ratios require integration over many more trials and are better suited to the gradual, implicit learning system of the basal ganglia. This interpretation is consistent with the hypothesis in the literature that cortical learning is impaired in patients, while striatal learning is intact (Gold et al 2008). We also examined win-stay/lose-shift behavior during early learning, and found that win-stay behavior (and, to a lesser extent, all stay behavior) was less frequent in patients than in controls. This suggests that patients relied on explicit strategy use to a lesser extent than controls, which is also consistent with the hypothesis that explicit mechanisms may have been impaired in this population.

While the whole-brain results from the full acquisition phase showed intact activation among patients in regions associated with cognitive control, there were some results that are consistent with the hypothesis of impaired cortical learning in this group. During the early learning phase, several regions involved in cognitive control demonstrated reduced overall choice-related activation in patients as compared to controls, which is consistent with a reduction in explicit learning during the early learning phase. Further, within the AB pair, B choices were associated with reduced activation in error- and conflict-processing regions including dorsal ACC, anterior prefrontal cortex, and thalamus, as well as OFC, medial temporal lobe, and cerebellar regions associated with executive control such as crus I and II (Habas et al 2009) among patients. In some of these regions, including bilateral anterior prefrontal cortex and left superior parietal lobule, this reduction in activation for B choices was associated with worse performance on the AB pair (and, for SPL, reduced win-stay behavior). Reduced error responses among individuals with schizophrenia are well established, and may reflect a failure to engage evaluative processes that contributes to impaired learning (Polli et al 2008). However, it is important to keep in mind that this relationship may simply represent larger error responses among those participants who are

doing well enough to know that “B” choices are less likely to yield reward, and may not be causally related to learning performance.

At the time of feedback, however, whole-brain activity was remarkably similar between patients and controls. Analyses of both positive versus negative feedback and prediction error analyses revealed patterns where reward regions responded to positive feedback or prediction errors and cognitive control regions responded to negative feedback or prediction error, a pattern that was intact in patients. There were few group differences at the time of feedback, and those we did see were small effects in regions not typically associated with any of the learning systems discussed here. One exception is VMPFC, which demonstrated effects of both choice type and feedback type that were absent in the patient group; however, these effects did not resemble a prediction error in controls and did not correlate with behavior or clinical symptoms. However, despite the intact activation at a group level, activation in response to positive feedback correlated with anhedonia/avolition in the patient group in both striatal and cortical regions. This finding is consistent with the hypothesis that deficits in responses to positive feedback contribute to these symptoms, although evidence for this interpretation is weakened by the fact that neither activation in these regions nor anhedonia/avolition severity correlated with gain learning rate.

While there is some evidence here that is suggestive that deficits in cortical learning mechanisms may have contributed to impaired performance in patients, the analyses of learning phase here were not individualized to each subject’s performance and may therefore have obscured important differences in these mechanisms, which may be active only at the earliest stages of learning. While the reinforcement learning model data was fit to individual subject behavior, these analyses may have been influenced by the fact that they examined brain activity during the full acquisition phase, which for many subjects included a long period after learning had plateaued. This design was chosen with gradual striatal learning in mind, and while it allows a preliminary look at cortical learning, additional analyses that give a more nuanced picture of cortical learning in this task are possible. In future work, we plan to examine pre-plateau learning as defined on a subject-by-subject basis in order to more accurately assess the contributions of cortical learning systems during early trials. We also plan to examine choice- and feedback-related activity associated with win-stay/lose-shift behavior to determine whether there are abnormalities in cortical activation associated with the alterations in strategy use in the patient group.

In addition to the limitations discussed above, the conclusions that can be drawn from this study about individuals with schizophrenia are limited by the fact that patients examined here were taking antipsychotic medications. Correlations with dose equivalents revealed increased NoGo learning in patients with higher medication doses, but no significant relationships between dose and brain activity. Interestingly, increased NoGo learning is what the Frank model predicts for greater levels of D2R antagonism, meaning that this relationship actually provides support for the model (Frank and O'Reilly R 2006; Waltz et al 2007). Perhaps surprisingly, studies examining striatal activation in schizophrenia tend to find reduced striatal activation for unmedicated patients, with intact activation for patients taking atypical antipsychotics, including some direct evidence of a normalizing effect of starting these medications (Nielsen et al 2012). The present study lends further support to these findings by demonstrating intact striatal activation in a population of patients primarily taking atypical antipsychotic medications.

In conclusion, this study demonstrated behavioral evidence of impaired learning from positive versus negative feedback, as well as impaired learning of stimuli with high versus low reinforcement ratios, among medicated patients with schizophrenia. Striatal activation was intact in the patient group at the time of choice and feedback, including intact prediction error activity. Cortical responses to feedback and prediction errors were also similar between groups. At the time of choice, patients failed to recruit cognitive control regions to the same extent as controls during early learning, and showed reduced responses in cognitive control and error-related regions when making low-probability choices, which correlated with performance. These findings are suggestive of alterations in cortical, but not basal ganglia, reinforcement learning mechanisms in the patient group. Severity of anhedonia and avolition in patients were not related to behavioral learning impairments or to prediction error activity, but were associated with reduced responses to positive feedback in caudate and bilateral DLPFC, suggesting a relationship between these symptoms and altered processing of positive feedback in patients. Future directions will further explore cortical activation during early learning and explicit strategy use to gain further insight into the contributions of explicit mechanisms to reinforcement learning in this population.

Supplementary Materials

Movement Data:

Table S1: Incremental (frame-by-frame) and absolute movement (root-mean-square values in mm), as well as mean voxelwise standard devation values for bolds 1-10

GROUP	BOLD	INCREMENTAL	ABSOLUTE	SD
CON	1	0.13 (0.07)	0.37 (0.27)	11.34 (2.83)
	2	0.18 (0.19)	0.38 (0.29)	11.83 (3.6)
	3	0.15 (0.09)	0.37 (0.26)	11.76 (2.74)
	4	0.17 (0.14)	0.43 (0.31)	12.5 (3.38)
	5	0.19 (0.15)	0.42 (0.28)	12.64 (3.42)
	6	0.19 (0.12)	0.57 (0.56)	13.69 (4.27)
	7	0.26 (0.29)	0.62 (0.59)	14.15 (5.21)
	8	0.23 (0.2)	0.54 (0.45)	13.77 (4.48)
	9	0.22 (0.23)	0.58 (0.41)	14.28 (4.45)
	10	0.24 (0.21)	0.58 (0.43)	14.14 (4.38)
SCZ	1	0.17 (0.16)	0.4 (0.46)	11.97 (4.16)
	2	0.18 (0.12)	0.38 (0.35)	12.03 (3.8)
	3	0.2 (0.22)	0.4 (0.36)	12.49 (4.83)
	4	0.18 (0.11)	0.47 (0.38)	13.08 (3.99)
	5	0.19 (0.12)	0.38 (0.31)	12.09 (3.16)
	6	0.21 (0.16)	0.5 (0.54)	13.64 (4.07)
	7	0.21 (0.2)	0.52 (0.63)	13.47 (4.37)
	8	0.24 (0.28)	0.49 (0.6)	13.44 (5.61)
	9	0.24 (0.3)	0.47 (0.45)	13.57 (4.97)
	10	0.2 (0.12)	0.4 (0.21)	12.91 (3.25)

Feedback ANOVA: Effects of Choice

Figure S1. Feedback ANOVA: Choice x Time Interactions

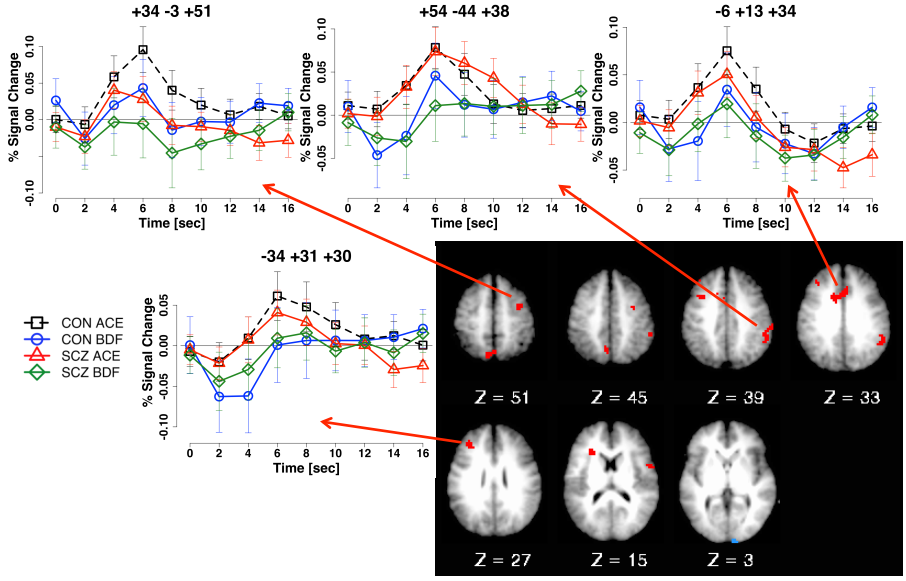


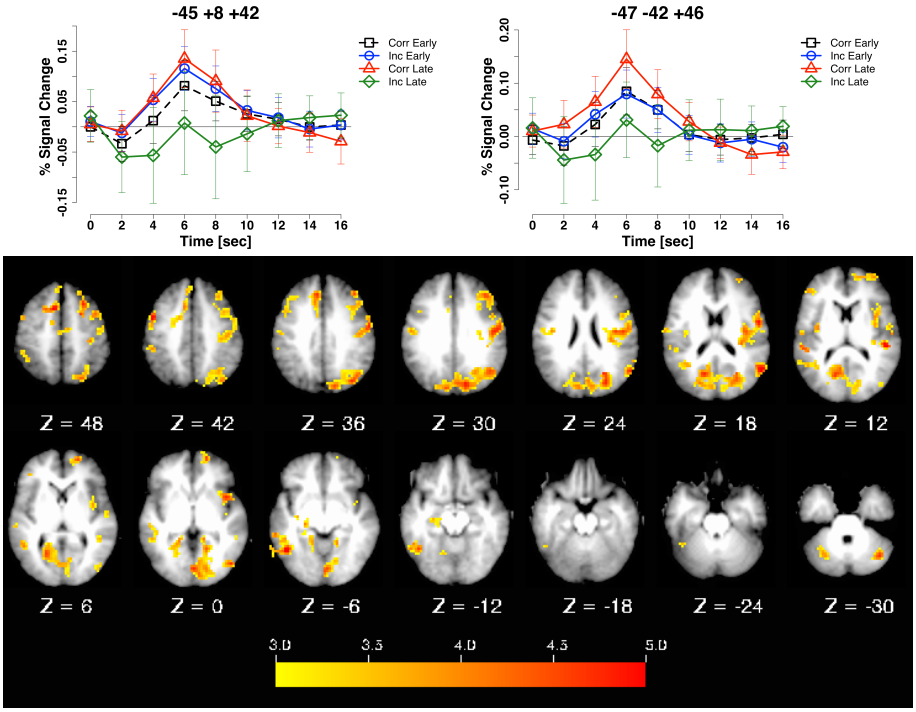
Table S2. Feedback ANOVA: Effects of choice

Effect	Region	BA	Talairach	Voxels	Z	Activation Pattern
Choice x Time	Middle Frontal Gyrus	9	-34_+31_+30	32	4.19	(+) ACE; (-) BDF
	ACC/pSMA	32	-6_+13_+34	52	4.18	ACE > BDF
	Anterior Insula		-26_+20_+16	14	4.18	ACE > BDF
	DLPFC	9	-35_+11_+37	17	4.02	ACE > BDF
	dPMC	6	+34_-3_+51	32	4.27	ACE > BDF
	Inferior Parietal Lobule	40	+54_-44_+38	67	4.84	ACE > BDF
	PPC	7	-6_-64_+48	34	4.31	ACE > BDF
	Precentral Gyrus	6	+53_+2_+16	15	4.45	ACE > BDF
	Cuneus	18	+10_-102_+1	16	4.39	BDF > ACE
Choice x FB x Time	Middle Frontal Gyrus	10	+44_+45_+15	23	4.29	BDF Pos > others

Table S3. Feedback ANOVA: Interactions between learning phase and choice

Region	BA	Talairach	Voxels	Z	Pattern	
					Early	Late
Inferior Parietal Lobule	40	-47_-42_+46	51	4.36	NS	ACE > BDF
Insula	13	+33_-22_+25	169	4.83	NS	ACE > BDF
Medial Frontal Gyrus	6	+3_-14_+48	21	3.92	NS	ACE > BDF
Middle Frontal Gyrus	8	-45_+8_+42	105	5.15	NS	ACE > BDF
Parietal White Matter	-	-45_-15_+24	72	4.84	NS	ACE > BDF
Postcentral Gyrus	43	+53_-12_+18	119	4.99	NS	ACE > BDF
Postcentral Gyrus	4	+42_-17_+45	115	4.19	NS	ACE > BDF
Precuneus	19	+16_-79_+38	113	4.24	NS	ACE > BDF
Superior Temporal Gyrus	22	-40_-53_+11	78	4.47	NS	ACE > BDF
Superior Temporal Gyrus	41	-50_-36_+15	39	4.86	NS	ACE > BDF
Cerebellar Tonsil	-	-32_-51_-46	72	4.81	BDF > ACE	ACE > BDF
Cerebellar Tonsil	-	+35_-62_-38	169	4.85	BDF > ACE	ACE > BDF
Claustrum	-	+38_+00_+07	113	4.51	BDF > ACE	ACE > BDF
Cuneus	18	+07_-77_+26	171	4.18	BDF > ACE	ACE > BDF
Cuneus	18	+25_-77_+21	171	4.32	BDF > ACE	ACE > BDF
Cuneus	18	-15_-80_+23	136	4.10	BDF > ACE	ACE > BDF
Inferior Frontal Gyrus	13	+33_+22_+09	26	3.85	BDF > ACE	ACE > BDF
Inferior Temporal Gyrus	20	-50_-56_-11	95	5.45	BDF > ACE	ACE > BDF
Lingual Gyrus	18	+09_-81_+02	211	4.42	BDF > ACE	ACE > BDF
Lingual Gyrus	19	+15_-51_-01	41	4.42	BDF > ACE	ACE > BDF
Lingual Gyrus	19	+23_-71_-01	39	3.83	BDF > ACE	ACE > BDF
Lingual Gyrus	19	-15_-46_-01	35	4.13	BDF > ACE	ACE > BDF
Middle Frontal Gyrus	6	+40_+06_+47	91	4.62	BDF > ACE	ACE > BDF
Middle Occipital Gyrus	19	+44_-78_+05	56	4.32	BDF > ACE	ACE > BDF
Posterior Cingulate	30	-19_-59_+05	93	4.60	BDF > ACE	ACE > BDF
Precentral Gyrus	9	+36_+24_+36	167	4.80	BDF > ACE	ACE > BDF
Sub-Gyral	37	-52_-40_-03	72	4.92	BDF > ACE	ACE > BDF
Superior Temporal Gyrus	41	+48_-38_+08	52	4.75	BDF > ACE	ACE > BDF
Tuber	-	-31_-61_-30	58	4.57	BDF > ACE	ACE > BDF
Inferior Frontal Gyrus	9	+49_+04_+23	61	4.24	NS	(+) ACE, (-) BDF
Medial Frontal Gyrus	6	+0_-22_+70	19	4.12	NS	(+) ACE, (-) BDF
Posterior Cingulate	31	-09_-65_+14	148	4.60	NS	(+) ACE, (-) BDF
Precentral Gyrus	6	+52_-05_+39	88	4.78	NS	(+) ACE, (-) BDF
Superior Frontal Gyrus	6	+25_+06_+57	133	4.95	NS	(+) ACE, (-) BDF
Superior Frontal Gyrus	6	-13_+14_+48	302	4.82	NS	(+) ACE, (-) BDF
Superior Temporal Gyrus	39	+56_-62_+19	95	5.02	(+) BDF Early, (-) ACE Early	(+) ACE, (-) BDF
Inferior Frontal Gyrus	46	-42_+34_+10	21	4.32	BDF > ACE	(+) ACE, (-) BDF
Medial Frontal Gyrus	10	+20_+51_+8	101	5.66	BDF > ACE	(+) ACE, (-) BDF
Precuneus	19	+34_-67_+39	234	4.44	BDF > ACE	(+) ACE, (-) BDF
Superior Frontal Gyrus	6	+11_+10_+66	14	4.01	BDF > ACE	(+) ACE, (-) BDF
Transverse Temporal Gyrus	42	-58_-15_+15	23	4.54	NS	BDF Shifted Later
Parahippocampal Gyrus	28	-23_-26_-8	49	4.49	BDF > ACE	NS

Figure S2: Feedback Learning Phase ANOVA: Learning Phase x Choice x Time



Chapter 5.

Conclusions and Future Directions

The preceding chapters examined the potential relationships between impairments in three candidate processes involved in translating reward information into goal-directed behavior and symptoms of anhedonia and avolition in schizophrenia: hedonics, reward prediction, and reinforcement learning. We chose to examine these processes because they are not only necessary for the generation of goal-directed behavior, but they are likely candidates for alteration in schizophrenia given their dependence on dopaminergic systems thought to be disrupted in this illness. However, a number of higher-level cognitive processes also contribute to forming goal-directed action plans, which may also be disrupted in patients experiencing these symptoms (Figure 5.1). Here, we discuss how our results speak to the relative contributions of these different mechanisms to motivational impairments in schizophrenia, as well as their implications for future work in this field.

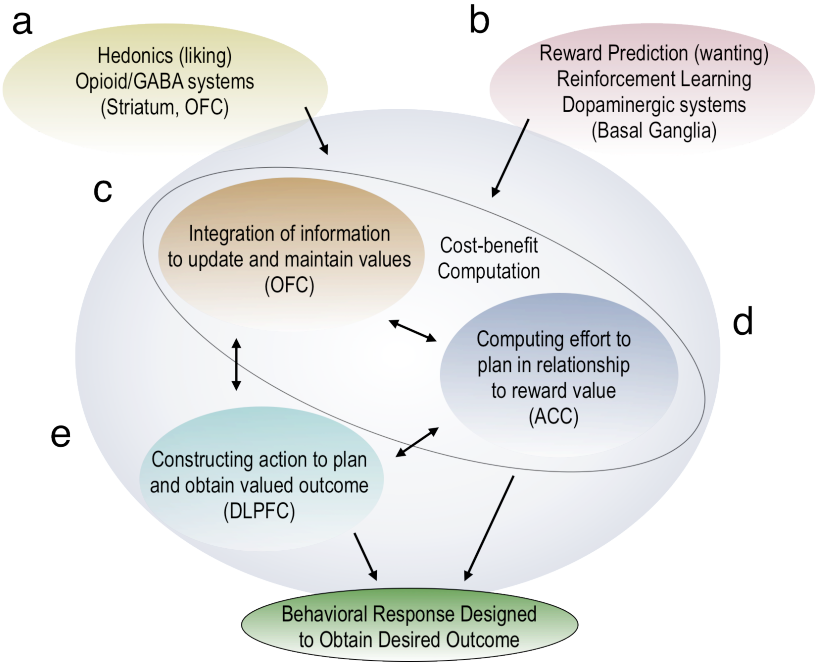


Figure 5.1. Components of reward-to-outcome translation. Figure adapted from (Barch and Dowd 2010). This figure outlines a conceptual overview of several processes thought to be involved in translating reward information into goal-directed behaviors, which may be compromised in schizophrenia.

Hedonics, reward prediction, and reinforcement learning: Summary and implications.

In Chapter 2, we examined hedonics (Figure 5.1a) in schizophrenia using self-reports and brain responses to emotionally evocative stimuli. We found that as a group, individuals with schizophrenia

demonstrated reduced valence, but not arousal, experiences of both positive and negative stimuli. However, we also found that individuals with more severe anhedonia reported experiencing less emotion in response to both positive and negative stimuli, a relationship that held among both patients and controls and fully mediated the differences between groups. These behavioral findings suggest that the relationship between reduced emotional experience and greater anhedonia are not specific to schizophrenia, but may reflect a relationship with the clinical dimension of anhedonia that crosses diagnostic boundaries. Further, they suggest that this relationship is not specific to a disruption in hedonic processes, because negative experiences were affected as well. This pattern of blunting for both positive and negative experiences has been reported in past studies in both patients with schizophrenia and controls, (Burbridge and Barch 2007; Herbener et al 2007) as well as in patients with major depressive disorder (Treadway and Zald 2011). Further studies of anhedonia in healthy adults as well as other psychiatric and neurological disorders may shed light on the extent to which this relationship is generalizable. In terms of fMRI results, this study demonstrated patterns of activation across emotional conditions that did not differ between groups at a whole-brain level, though ROI analysis revealed regions in left putamen and right ventral striatum with altered activation for positive stimuli in patients. In addition, increased anhedonia was associated with reduced activation for positive vs. negative stimuli in bilateral amygdala and right ventral striatum in patients, and in bilateral caudate in controls. These results are consistent with the possibility that hedonic deficits are associated with impairments in striatal and amygdala processing of pleasant stimuli, but due to nonspecificity in the anhedonia rating scales, may reflect motivational deficits as well.

In Chapter 3, we examined reward prediction in schizophrenia using a Pavlovian paradigm with monetary reward. We found intact striatal responses to reward receipt in the patient group that were not sensitive to anhedonia/avolition severity. These findings are somewhat at odds with the results of Chapter 2, which show abnormal responses to pleasant experiences among patients, particularly those higher in anhedonia. However, they do agree with several studies in the literature in showing intact activation in response to receipt of monetary reward (Kirsch et al 2007; Schlagenhauf et al 2009; Simon et al 2010; Walter et al 2009). One potential explanation for this is that well-predicted rewards are expected to elicit a minimal dopaminergic response, and tasks examining pleasant events that are not preceded by predictive

cues, or that deliberately elicit prediction errors, may be more sensitive to alterations in feedback responses among individuals with schizophrenia, at least those with higher self-reported anhedonia. It is also possible that anhedonic individuals are able to respond normally to simple monetary rewards, but have more difficulty responding to the complex scenes and faces used in emotion studies. This study also found that during reward anticipation, greater anhedonia/avolition was associated with smaller responses in ventral striatum in patients, consistent with the hypothesis that reductions in striatal reward prediction contribute to anhedonia in schizophrenia. In addition, the same relationship was seen in VMPFC among both patients and controls, again suggesting that there may be aspects of the relationship between anhedonia and reward processing that cross diagnostic boundaries.

Chapter 4 examined probabilistic reinforcement learning in schizophrenia and found behavioral evidence for impairments in learning from positive feedback in patients as compared to controls. We found no evidence of altered striatal responses to choice, feedback, or prediction error, but did find evidence suggestive of impairments in cortical learning systems relying on cognitive control regions including DLPFC. Reduced responses to positive feedback in caudate and posterior DLPFC were associated with greater anhedonia/avolition severity among patients, but not controls. The finding of a relationship between activation for positive feedback and motivational deficits in the reinforcement learning study, but not in the reward-prediction study (or others in the literature), suggests that deficits in feedback responses among amotivated individuals may be related to using feedback information to update reward representations and guide future choices. The fact that we found this relationship in bilateral DLPFC lends further support to this hypothesis, as maintaining a dynamic representation of the state of the organism and the environment is a function thought to be supported by this region. Specifically, posterior DLPFC has been implicated in decision-making, and has high connectivity with posterior parietal cortex (Cieslik et al 2012). Posterior parietal cortex has been proposed to act as a salience-specific behavioral integrator, incorporating information about expected rewards and goal context with visuospatial, motor, and memory information under top-down influence from DLPFC (Gottlieb 2007). The reduced activation in response to positive feedback in posterior DLPFC regions among patients high in avolition may therefore represent disruption in a network responsible for integrating reward information with other contextual information required to guide appropriate behavior.

Taken together, a few commonalities stand out across these studies. First, all three studies examined populations of controls and patients who differed significantly in their levels of self-reported anhedonia/avolition. Second, all three studies identified activation patterns in reward processing systems that were surprisingly intact among the patient group. While there were some group differences that could be meaningful, in general these differences were either of small effect sizes, or were not found in regions associated with reward processing. Third, despite the relative lack of group differences, all three studies identified relationships between altered activation and severity of self-reported anhedonia/avolition among patients. This pattern of findings suggests that the processes examined here vary with individual differences in anhedonia/avolition in a meaningful way, but that there is additional variance in these symptoms that is unaccounted for by any of the three processes studied. One possibility, consistent with the cortical learning impairments and association of altered posterior DLPFC activity with anhedonia/avolition shown in Chapter 4, is that impairments in cortically mediated processes involved in the generation of goal-directed behavior also contribute to these symptoms.

From reward prediction to goal-directed behavior: cortical contributions

In the studies described here, we examined the hypothesis that reductions in goal-directed behavior result from disruptions in dopamine-mediated striatal reward valuation and prediction systems, which would reduce the ability of appetitive cues to drive behavior. However, these disruptions may be exacerbated by impairments in higher-level cognitive processes required to translate information about reward values and predictions into goal-directed behavior. These processes are illustrated in Figure 5.1 ((Barch and Dowd 2010), adapted from (Wallis 2007)). First, information from the hedonic, reward prediction, and reinforcement learning processes studied here must be integrated into representations of value. A prominent hypothesis posits that this computation is supported by OFC, which integrates the reinforcing properties of the stimulus with the internal state of the organism, as well as updating changes in reward contingencies (Wallis 2007). In addition, action plans must be generated using current information about values and goals, a process that may be supported by DLPFC (Miller and Cohen 2001). Further, candidate action plans must undergo cost-benefit analysis in which the effort associated with the plan is computed, purportedly by dorsal ACC (Croxson et al 2009; Salamone et al 2007), and

weighed against the expected value of potential rewards. Interaction between these processes allows selection of the action plan most likely to yield the desired outcome, which is then translated into a behavioral response. Below, we discuss the existing evidence and need for future work to determine whether disruptions in these processes also contribute to motivational deficits in schizophrenia.

Value computation and OFC function in schizophrenia: The hypothesis that motivational deficits in schizophrenia may reflect deficits in OFC-based value representation has been raised by Gold and colleagues, who have found behavioral evidence that individuals with schizophrenia have difficulty integrating and updating information about rewards and punishments (Gold et al 2008). Experimental paradigms probing OFC function, including probabilistic reversal learning and the Iowa Gambling Task, have both yielded fairly consistent evidence of impairment in schizophrenia (Ceaser et al 2008. Shurman, 2005 #158; Elliott et al 1995; Kester et al 2006; Kim et al 2009; Lee and Seo 2007; Martino et al 2007; Oades 1997; Pantelis et al 1999; Premkumar et al 2008; Sevy et al 2007; Turnbull et al 2006; Tyson et al 2004; Waltz and Gold 2007; Yip et al 2009). There is also evidence for an association between reduced OFC volume and negative symptoms in schizophrenia (Baare et al 1999; Gur et al 2000). Further, a study using a computational reinforcement learning model was able to reproduce the behavior of patients with high negative symptoms when deficits in putatively OFC-based functions were simulated (Gold et al 2012). However, as of yet, there is no neuroimaging data on OFC function in relationship to value representation in schizophrenia using tasks such as probabilistic reversal learning, making this a promising focus for future studies.

Effort computation and ACC function in schizophrenia: To our knowledge, there is no work addressing effort computations in schizophrenia. ACC dysfunction has been reported in schizophrenia in studies showing reduced error-related responses (Polli et al 2008) and reduced conflict-related activation (Kerns et al 2005). It is possible that these functions share common mechanisms with effort computation, as greater conflict and error are likely associated with a requirement for greater effort, but more work is needed to address this possibility.

Goal-directed action and DLPFC function in schizophrenia: There is a large body of evidence for impairments in cognitive functions thought to be mediated by DLPFC in schizophrenia (Barch 2005; Heinrichs and Zakzanis 1998; Lee and Park 2005), including working memory, context representation,

goal maintenance, and planning. Furthermore, there is robust evidence for altered DLPFC function in schizophrenia during cognitive control tasks (Minzenberg et al 2009). Given these findings, one appealing possibility is that a core deficit in the ability to represent goal and context information in working memory may underlie both cognitive and motivational deficits in schizophrenia (Figure 5.1e). Specifically, we speculate that these deficits may reflect impairments in the ability to maintain and utilize internal representations of emotional experiences, previous rewards, and motivational goals in order to drive goal-directed behavior. This hypothesis is consistent with the correlation between reduced feedback-related activation in posterior DLPFC and anhedonia/avolition severity in patients as discussed above. Further, the ability to use internal goal and context representations may be closely related to the prefrontal explicit learning mechanisms discussed in Chapter 4, which showed some evidence of impairment in schizophrenia.

One important direction for future research is to examine whether individuals with schizophrenia can adjust their cognitive performance under different motivational conditions, a process that may occur via modulation of DLPFC activity (Locke and Braver 2008). Several studies suggest that patients are not able to improve their performance on cognitive control tasks when offered monetary incentives (Green et al 1992; Hellman et al 1998; Roiser et al 2009; Vollema and van den Bosch 1995), though no neuroimaging studies have yet been conducted to examine the neural correlates of this impairment. In a different approach, recent work by Ursu and colleagues demonstrated reductions in DLPFC activation in patients during maintenance of affective information, which were correlated with negative symptom severity (Ursu et al 2011). While this relationship was shown for negative stimuli, the general pattern is consistent with the hypothesis that individuals with schizophrenia may have difficulty representing information about rewards and incentives that can be used to drive goal directed behavior, and provides a model for future work.

Anhedonia vs. Avolition: self-report limitations and future possibilities

Each of the studies conducted here relied on self-report measures to assess anhedonia and avolition severity. As discussed previously, this assessment strategy has two major limitations: first, that

such measures do not adequately separate the constructs of anhedonia and avolition, and second, that it is not clear whether patients are able to give veridical self-reports of remembered or anticipated pleasure, as opposed to pleasure experienced in-the-moment (Strauss and Gold 2012). The problem of measuring hedonic and motivational deficits separately is complex, given that motivational deficits may result from anhedonia, and amotivated individuals are likely to experience fewer pleasant events, decreasing overall pleasure experience (which may be misinterpreted as anhedonia). This close interrelation is likely why these two symptoms load onto a single factor in clinician-rated scales of negative symptoms such as the SANS and the BNSS (Messinger et al 2011). Further complicating this distinction is the fact that most scales assessing "anhedonia" predate the knowledge of dissociable hedonic and motivational neural systems, and tend to conflate these constructs. While we have tended to discuss a single dimension of anhedonia/avolition here for these reasons, we did explore the possibility of separating these constructs in each of the studies above. In Chapter 2, we found that the relationships with anhedonia were also present when using the SANS avolition scale. In Chapter 4, we found that correlations between self-report scales did not respect the boundaries of whether the scale was intended to assess anhedonia or avolition. Further, the correlations between anhedonia/avolition summary score and DLPFC activation in this study were present among several scales separately, including some newer scales intended to assess hedonics (e.g. TEPS consummatory pleasure) and some intended to assess motivation (e.g. the Apathy Scale). New scales such as the Clinical Assessment Interview for Negative Symptoms (CAINS) (Horan et al 2011) may help to disentangle to what extent these difficulties are due to scale nonspecificity, and to what extent they reflect clinical co-occurrence of anhedonia and avolition. The CAINS is unique among symptom scales in that it was developed using an iterative, data-driven approach to item generation, selection, and retention, and has undergone comprehensive assessment of its psychometric properties and validity in a large, diverse sample of patients. While anhedonia and avolition load onto a single factor in this scale as well, the items addressing these symptoms carefully distinguish between the frequency of pleasant or goal-directed events and the internal experience when these events occur, information that may potentially be used to tease apart effects of hedonic versus motivational deficits.

The second problem with self-report measures is that scores on these measures may be influenced by cognitive deficits that make reporting remembered or anticipated pleasure difficult in

patients. This possibility is consistent with our findings of group differences in self-reported anhedonia/avolition, but few group differences in brain activation on tasks that probe the reward system in response to stimuli presented in-the-moment. However, if anhedonia scores in patients are driven primarily by difficulties in reporting, rather than experiencing, pleasure, we would not expect them to be associated with in-the-moment emotional experience, as we found in Chapter 2. Further, we would not expect them to show the same relationships with behavioral or neuroimaging variables that were present in controls, as we found in Chapters 2 and 3. We feel that these results provide evidence that the self-reports used here captured meaningful variation in anhedonia/avolition severity over and above any influence of cognitive deficits in patients, though it does not rule out some contribution of cognitive impairments.

Given these concerns, future studies may benefit from obtaining anhedonia and avolition measures that do not rely on self-report, which will also provide independent estimates of hedonic versus motivational deficits. Assessments of avolition can be obtained from informants such as family members who can provide accurate information about a patient's day-to-day activities. Accurate measurements of anhedonia may be best obtained using laboratory-based assessments of in-the-moment consummatory behaviors.

Conclusions

In conclusion, the studies included here provide evidence that impairments in reward processing functions including hedonics, reward prediction, and reinforcement learning are likely contributors, but are not sufficient to explain, anhedonia and avolition in medicated chronic outpatients with schizophrenia. We found evidence that reductions in both positive and negative emotional experience are related to increased anhedonia, and that anhedonia was associated with reduced ventral striatal activation for positive versus negative stimuli. We also found that anhedonia/avolition was related to reduced anticipatory activity in ventral striatum and VMPFC during reward prediction, and to reduced positive feedback responses in caudate and DLPFC during reinforcement learning. However, group differences in functional activation in these studies, particularly in striatal regions, were few, and do not seem likely to

fully explain the much larger group differences in anhedonia/avolition severity. We therefore hypothesize that deficits in additional processes required to translate information about received or anticipated rewards into goal-directed behavior also contribute to motivational deficits in schizophrenia. In particular, we speculate that deficits in reward prediction and learning may be exacerbated by the inability to integrate, maintain, and flexibly update information about rewards, goals, and context in order to form action plans. What is needed next is research that tests the ability to use and maintain internal representations of reward information to modulate behavior and brain function in schizophrenia, along with work that characterizes links between deficits in this ability and everyday function in this illness.

References

Abrams DJ, Rojas DC, Arciniegas DB (2008): Is schizoaffective disorder a distinct categorical diagnosis? A critical review of the literature. *Neuropsychiatric disease and treatment* 4:1089-1109.

Ahsan RL, Allom R, Gousias IS, Habib H, Turkheimer FE, Free S, et al (2007): Volumes, spatial extents and a probabilistic atlas of the human basal ganglia and thalamus. *Neuroimage* 38:261-270.

Aleman A, Kahn RS (2005): Strange feelings: do amygdala abnormalities dysregulate the emotional brain in schizophrenia? *Prog Neurobiol* 77:283-298.

Andreasen NC (1983a): The scale for the assessment of negative symptoms (SANS). Iowa City: University of Iowa.

Andreasen NC (1983b): The scale for the assessment of positive symptoms (SAPS). Iowa City: University of Iowa.

Andreasen NC, Arndt S, Alliger R, Miller D, Flaum M (1995): Symptoms of schizophrenia: Methods, meanings, and mechanisms. *Arch Gen Psychiatry* 52:341-351.

Baare WFC, Pol HEH, Hijman R, Mali WPT, Viergever MA, Kahn RS (1999): Volumetric analysis of frontal lobe regions in schizophrenia: Relation to cognitive function and symptomatology. *Biol Psychiatry* 45:1597-1605.

Bandettini PA, Jesmanowiz J, Wong EC, Hyde JS (1993): Processing strategies for time-course data sets in functional MRI of the human brain. *Magn Reson Med* 30:161-173.

Barch DM (2005): The Cognitive Neuroscience of Schizophrenia. *Annu Rev Clin Psychol* 1:321-353.

Barch DM, Csernansky J, Conturo T, Snyder AZ, Ollinger J (2002): Working and long-term memory deficits in schizophrenia. Is there a common underlying prefrontal mechanism? *J Abnorm Psychol* 111:478-494.

Barch DM, Dowd EC (2010): Goal representations and motivational drive in schizophrenia: the role of prefrontal-striatal interactions. *Schizophr Bull* 36:919-934.

Barrett LF, Mesquita B, Ochsner KN, Gross JJ (2007): The experience of emotion. *Annu Rev Psychol* 58:373-403.

Beck SM, Locke HS, Savine AC, Jimura K, Braver TS (2010): Primary and secondary rewards differentially modulate neural activity dynamics during working memory. *PLoS ONE* 5:e9251.

Beninger RJ, Wasserman J, Zanibbi K, Charbonneau D, Mangels J, Beninger BV (2003): Typical and atypical antipsychotic medications differentially affect two nondeclarative memory tasks in schizophrenic patients: a double dissociation. *Schizophr Res* 61:281-292.

Benjamini Y, Hochberg Y (1995): Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society* 57:289-300.

Berridge KC (2004): Motivation concepts in behavioral neuroscience. *Physiol Behav* 81:179-209.

Berridge KC (2007): The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology (Berl)* 191:391-431.

Berridge KC, Kringelbach ML (2008): Affective neuroscience of pleasure: reward in humans and animals. *Psychopharmacology (Berl)* 199:457-480.

Berridge KC, Robinson TE (1998): What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res Brain Res Rev* 28:309-369.

Blanchard JJ, Bellack AS, Mueser KT (1994): Affective and social-behavioral correlates of physical and social anhedonia in schizophrenia. *J Abnorm Psychol* 103:719-728.

Bleuler E (1911): *Dementia Praecox, or the Group of Schizophrenias*. New York: International Universities Press.

Bradley MM, Lang PJ (1999): Affective norms for English words (ANEW): Instruction manual and affective ratings. Gainesville: University of Florida.

Burbridge JA, Barch DM (2007): Anhedonia and the experience of emotion in individuals with schizophrenia. *J Abnorm Psychol* 116:30-42.
Burgdorf J, Panksepp J (2006): The neurobiology of positive emotions. *Neurosci Biobehav Rev* 30:173-187.
Calabresi P, Pisani A, Centonze D, Bernardi G (1997): Synaptic plasticity and physiological interactions between dopamine and glutamate in the striatum. *Neurosci Biobehav Rev* 21:519-523.
Ceaser AE, Goldberg TE, Egan MF, McMahon RP, Weinberger DR, Gold JM (2008): Set-shifting ability and schizophrenia: a marker of clinical illness or an intermediate phenotype? *Biol Psychiatry* 64:782-788.
Chapman LJ, Chapman JP (1978): The revised physical anhedonia scale: University of Wisconsin: Madison.
Chapman LJ, Chapman JP, Raulin ML (1976): Scales for physical and social anhedonia. *J Abnorm Psychol* 85:374-382.
Cheng GL, Tang JC, Li FW, Lau EY, Lee TM (2012): Schizophrenia and risk-taking: impaired reward but preserved punishment processing. *Schizophr Res* 136:122-127.
Cheniaux E, Landeira-Fernandez J, Lessa Telles L, Lessa JL, Dias A, Duncan T, et al (2008): Does schizoaffective disorder really exist? A systematic review of the studies that compared schizoaffective disorder with schizophrenia or mood disorders. *J Affect Disord* 106:209-217.
Cieslik EC, Zilles K, Caspers S, Roski C, Kellermann TS, Jakobs O, et al (2012): Is There "One" DLPFC in Cognitive Action Control? Evidence for Heterogeneity From Co-Activation-Based Parcellation. *Cereb Cortex*.
Cohen AS, Minor KS (2008): Emotional Experience in Patients With Schizophrenia Revisited: Meta-analysis of Laboratory Studies. *Schizophr Bull*.
Cohen MX, Frank MJ (2009): Neurocomputational models of basal ganglia function in learning, memory and choice. *Behav Brain Res* 199:141-156.
Cole MW, Schneider W (2007): The cognitive control network: Integrated cortical regions with dissociable functions. *Neuroimage* 37:343-360.
Corlett PR, Murray GK, Honey GD, Aitken MR, Shanks DR, Robbins TW, et al (2007): Disrupted prediction-error signal in psychosis: evidence for an associative account of delusions. *Brain* 130:2387-2400.
Crespo-Facorro B, Paradiso S, Andreasen NC, O'Leary DS, Watkins GL, Ponto LL, et al (2001): Neural mechanisms of anhedonia in schizophrenia: a PET study of response to unpleasant and pleasant odors. *JAMA* 286:427-435.
Croxson PL, Walton ME, O'Reilly JX, Behrens TE, Rushworth MF (2009): Effort-based cost-benefit valuation and the human brain. *J Neurosci* 29:4531-4541.
Curtis CE, Lebow B, Lake DS, Katsanis J, Iacono WG (1999): Acoustic startle reflex in schizophrenic patients and their first-degree relatives: Evidence of normal emotional modulation. *Psychophysiology* 36:469-475.
Dayan P, Balleine BW (2002): Reward, motivation, and reinforcement learning. *Neuron* 36:285-298.
Dosenbach NU, Visscher KM, Palmer ED, Miezin FM, Wenger KK, Kang HC, et al (2006): A core system for the implementation of task sets. *Neuron* 50:799-812.
Dowd EC, Barch DM (2010): Anhedonia and emotional experience in schizophrenia: neural and behavioral indicators. *Biol Psychiatry* 67:902-911.
Doya K (1999): What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Netw* 12:961-974.
Eckblad M, Chapman LJ, Chapman JP, Mishlove M (1982): The revised social anhedonia scale: University of Wisconsin: Madison.
Ekman P FW (1976): *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

Elliott R, McKenna PJ, Robbins TW, Sahakian BJ (1995): Neuropsychological evidence for frontostriatal dysfunction in schizophrenia. *Psychol Med* 25:619-630.
Epstein J, Pan H, Kocsis JH, Yang Y, Butler T, Chusid J, et al (2006): Lack of ventral striatal response to positive stimuli in depressed versus normal subjects. *Am J Psychiatry* 163:1784-1790.
Erlenmeyer-Kimling L, Cornblatt BA, Rock D, Roberts S, Bell M, West A (1993): The New York High-Risk Project: anhedonia, attentional deviance, and psychopathology. *Schizophr Bull* 19:141-153.
First M, Spitzer R, Gibbon M, Williams J (2001): *Structured clinical interview for DSM-IV-TR axis I disorders-patient edition*: Biometrics Research Department, New York State Psychiatric Institute.
Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995): Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magn Reson Med* 33:636-647.
Frank M, Claus ED (2006): Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychol Rev* 113:300-326.
Frank M, Seeberger LC, O'reilly RC (2004): By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 306:1940-1943.
Frank MJ, Moustafa AA, Haughey HM, Curran T, Hutchison KE (2007): Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proc Natl Acad Sci U S A* 104:16311-16316.
Frank MJ, O'Reilly R C (2006): A mechanistic account of striatal dopamine function in human cognition: psychopharmacological studies with cabergoline and haloperidol. *Behav Neurosci* 120:497-517.
Friston KJ, Holmes AP, Worsley KJ, Poline J-P, Frith CD, Frackowiak RSJ (1995): Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp* 2:189-210.
Friston KJ, Jezzard P, Turner R (1994): The analysis of functional MRI time series. *Hum Brain Mapp* 2:1-45.
Fusar-Poli P, Meyer-Lindenberg A (2012): Striatal Presynaptic Dopamine in Schizophrenia, Part II: Meta-Analysis of [18F/11C]-DOPA PET Studies. *Schizophr Bull*.
Gard DE, Kring AM, Gard MG, Horan WP, Green MF (2007): Anhedonia in schizophrenia: distinctions between anticipatory and consummatory pleasure. *Schizophr Res* 93:253-260.
Gardner DM, Murphy AL, O'Donnell H, Centorrino F, Baldessarini RJ (2010): International consensus study of antipsychotic dosing. *Am J Psychiatry* 167:686-693.
Gold JM, Waltz JA, Matveeva TM, Kasanova Z, Strauss GP, Herbener ES, et al (2012): Negative symptoms and the failure to represent the expected reward value of actions: behavioral and computational modeling evidence. *Arch Gen Psychiatry* 69:129-138.
Gold JM, Waltz JA, Prentice KJ, Morris SE, Heerey EA (2008): Reward processing in schizophrenia: a deficit in the representation of value. *Schizophr Bull* 34:835-847.
Goto Y, Grace AA (2007): The Dopamine System and the Pathophysiology of Schizophrenia: A Basic Science Perspective. *Int Rev Neurobiol* 78C:41-68.
Gottlieb J (2007): From thought to action: the parietal cortex as a bridge between perception, action, and cognition. *Neuron* 53:9-16.
Gradin VB, Kumar P, Waiter G, Ahearn T, Stickle C, Milders M, et al (2011): Expected value and prediction error abnormalities in depression and schizophrenia. *Brain* 134:1751-1764.
Green MF, Satz P, Ganzell S, Vaclav JF (1992): Wisconsin Card Sorting Test performance in schizophrenia: remediation of a stubborn deficit. *Am J Psychiatry* 149:62-67.
Grimm O, Vollstadt-Klein S, Krebs L, Zink M, Smolka MN (2012): Reduced striatal activation during reward anticipation due to appetite-provoking cues in chronic schizophrenia: a fMRI study. *Schizophr Res* 134:151-157.
Guillin O, Abi-Dargham A, Laruelle M (2007): Neurobiology of dopamine in schizophrenia. *Int Rev Neurobiol* 78:1-39.

Gur RC, Sara R, Hagendoorn M, Marom O, Hughett P, Macy L, et al (2002a): A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies. *J Neurosci Methods* 115:137-143.

Gur RE, Cowell PE, Latshaw A, Turetsky BI, Grossman RI, Arnold SE, et al (2000): Reduced dorsal and orbital prefrontal gray matter volumes in schizophrenia. *Arch Gen Psychiatry* 57:761-768.

Gur RE, McGrath C, Chan RM, Schroeder L, Turner T, Turetsky BI, et al (2002b): An fMRI study of facial emotion processing in patients with schizophrenia. *Am J Psychiatry* 159:1992-1999.

Habas C, Kamdar N, Nguyen D, Prater K, Beckmann CF, Menon V, et al (2009): Distinct cerebellar contributions to intrinsic connectivity networks. *J Neurosci* 29:8586-8594.

Habel U, Klein M, Shah NJ, Toni I, Zilles K, Falkai P, et al (2004): Genetic load on amygdala hypofunction during sadness in nonaffected brothers of schizophrenia patients. *Am J Psychiatry* 161:1806-1813.

Hall J, Harris JM, McKirdy JW, Johnstone EC, Lawrie SM (2007): Emotional memory in schizophrenia. *Neuropsychologia* 45:1152-1159.

Hall J, Whalley HC, McKirdy JW, Romaniuk L, McGonigle D, McIntosh AM, et al (2008): Overactivation of fear systems to neutral faces in schizophrenia. *Biol Psychiatry* 64:70-73.

Harvey PO, Armony J, Malla A, Lepage M (2010): Functional neural substrates of self-reported physical anhedonia in non-clinical individuals and in patients with schizophrenia. *J Psychiatr Res* 44:707-716.

Heerey EA, Bell-Warren KR, Gold JM (2008): Decision-making impairments in the context of intact reward sensitivity in schizophrenia. *Biol Psychiatry* 64:62-69.

Heerey EA, Gold JM (2007): Patients with schizophrenia demonstrate dissociation between affective experience and motivated behavior. *J Abnorm Psychol* 116:268-278.

Heinrichs RW, Zakzanis KK (1998): Neurocognitive deficit in schizophrenia: A quantitative review of the evidence. *Neuropsychology* 12:426-445.

Heinz A, Schlagenhauf F (2010): Dopaminergic dysfunction in schizophrenia: salience attribution revisited. *Schizophr Bull* 36:472-485.

Hellman SG, Kern RS, Neilson LM, Green MF (1998): Monetary reinforcement and Wisconsin Card Sorting performance in schizophrenia: why show me the money? *Schizophr Res* 34:67-75.

Hempel A, Hempel E, Schonknecht P, Stippich C, Schroder J (2003): Impairment in basal limbic function in schizophrenia during affect recognition. *Psychiatry Res* 122:115-124.

Herbener ES, Harrow M, Hill SK (2005): Change in the relationship between anhedonia and functional deficits over a 20-year period in individuals with schizophrenia. *Schizophr Res* 75:97-105.

Herbener ES, Rosen C, Khine T, Sweeney JA (2007): Failure of positive but not negative emotional valence to enhance memory in schizophrenia. *J Abnorm Psychol* 116:43-55.

Holt DJ, Kunkel L, Weiss AP, Goff DC, Wright CI, Shin LM, et al (2006): Increased medial temporal lobe activation during the passive viewing of emotional and neutral facial expressions in schizophrenia. *Schizophr Res* 82:153-162.

Horan W (2005): Anhedonia in Schizophrenia: A Review of Assessment Strategies. *Schizophr Bull* 32:259-273.

Horan WP, Green MF, Kring AM, Nuechterlein KH (2006): Does anhedonia in schizophrenia reflect faulty memory for subjectively experienced emotions? *J Abnorm Psychol* 115:496-508.

Horan WP, Kring AM, Gur RE, Reise SP, Blanchard JJ (2011): Development and psychometric validation of the Clinical Assessment Interview for Negative Symptoms (CAINS). *Schizophr Res* 132:140-145.

Howes OD, Egerton A, Allan V, McGuire P, Stokes P, Kapur S (2009): Mechanisms underlying psychosis and antipsychotic treatment response in schizophrenia: insights from PET and SPECT imaging. *Curr Pharm Des* 15:2550-2559.

Howes OD, Kambeitz J, Kim E, Stahl D, Slifstein M, Abi-Dargham A, et al (2012): The Nature of Dopamine Dysfunction in Schizophrenia and What This Means for Treatment: Meta-analysis of Imaging Studies. *Arch Gen Psychiatry*.

Hutton SB, Puri BK, Duncan LJ, Robbins TW, Barnes TR, Joyce EM (1998): Executive function in first-epsiode schizophrenia. *Psychol Med* 28:463-473.

Jazbec S, Pantelis C, Robbins T, Weickert T, Weinberger DR, Goldberg TE (2007): Intra-dimensional/extra-dimensional set-shifting performance in schizophrenia: impact of distractors. *Schizophr Res* 89:339-349.

Jensen J, Smith AJ, Willeit M, Crawley AP, Mikulis DJ, Vitcu I, et al (2007): Separate brain regions code for salience vs. valence during reward prediction in humans. *Hum Brain Mapp* 28:294-302.

Jensen J, Willeit M, Zipursky R, Savina I, Smith A, Menon M, et al (2008): The Formation of Abnormal Associations in Schizophrenia: Neural and Behavioral Evidence. *Neuropsychopharmacology* 33:473-479.

Johnston PJ, Stojanov W, Devir H, Schall U (2005): Functional MRI of facial emotion recognition deficits in schizophrenia and their electrophysiological correlates. *Eur J Neurosci* 22:1221-1232.

Joyce E, Hutton S, Mutsatsa S, Gibbins H, Webb E, Paul S, et al (2002): Executive dysfunction in first-episode schizophrenia and relationship to duration of untreated psychosis: the West London Study. *Br J Psychiatry Suppl* 43:s38-44.

Juckel G, Friedel E, Koslowski M, Witthaus H, Ozgurdal S, Gudlowski Y, et al (2012): Ventral striatal activation during reward processing in subjects with ultra-high risk for schizophrenia. *Neuropsychobiology* 66:50-56.

Juckel G, Schlagenhauf F, Koslowski M, Filonov D, Wüstenberg T, Villringer A, et al (2006a): Dysfunction of ventral striatal reward prediction in schizophrenic patients treated with typical, not atypical, neuroleptics. *Psychopharmacology (Berl)* 187:222-228.

Juckel G, Schlagenhauf F, Koslowski M, Wüstenberg T, Villringer A, Knutson B, et al (2006b): Dysfunction of ventral striatal reward prediction in schizophrenia. *Neuroimage* 29:409-416.

Kable JW, Glimcher PW (2007): The neural correlates of subjective value during intertemporal choice. *Nat Neurosci* 10:1625-1633.

Kapur S (2003): Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am J Psychiatry* 160:13-23.

Keedwell P, Andrew C, Williams S, Brammer M, Phillips M (2005): The neural correlates of anhedonia in Major Depressive Disorder. *Biol Psychiatry* 58:843-853.

Kegeles LS, Abi-Dargham A, Frankle WG, Gil R, Cooper TB, Slifstein M, et al (2010): Increased synaptic dopamine function in associative regions of the striatum in schizophrenia. *Arch Gen Psychiatry* 67:231-239.

Keri S, Kelemen O, Szekeres G, Bagoczky N, Erdelyi R, Antal A, et al (2000): Schizophrenics know more than they can tell: probabilistic classification learning in schizophrenia. *Psychol Med* 30:149-155.

Keri S, Nagy O, Kelemen O, Myers CE, Gluck MA (2005): Dissociation between medial temporal lobe and basal ganglia memory systems in schizophrenia. *Schizophr Res* 77:321-328.

Kerns JG, Cohen JD, MacDonald AW, 3rd, Johnson MK, Stenger VA, Aizenstein H, et al (2005): Decreased conflict- and error-related activity in the anterior cingulate cortex in subjects with schizophrenia. *Am J Psychiatry* 162:1833-1839.

Kester HM, Sevy S, Yechiam E, Burdick KE, Cervellione KL, Kumra S (2006): Decision-making impairments in adolescents with early-onset schizophrenia. *Schizophr Res* 85:113-123.

Kim YT, Lee KU, Lee SJ (2009): Deficit in decision-making in chronic, stable schizophrenia: from a reward and punishment perspective. *Psychiatry Investig* 6:26-33.

Kirkpatrick B, Fenton WS, Carpenter WT, Jr., Marder SR (2006): The NIMH-MATRICS consensus statement on negative symptoms. *Schizophr Bull* 32:214-219.

Kirkpatrick B, Strauss GP, Nguyen L, Fischer BA, Daniel DG, Cienfuegos A, et al (2011): The brief negative symptom scale: psychometric properties. *Schizophr Bull* 37:300-305.
Kirsch P, Ronshausen S, Mier D, Gallhofer B (2007): The influence of antipsychotic treatment on brain reward system reactivity in schizophrenia patients. *Pharmacopsychiatry* 40:196-198.
Knapp M, Mangalore R, Simon J (2004): The global costs of schizophrenia. *Schizophr Bull* 30:279-293.
Knutson B, Adams CM, Fong GW, Hommer D (2001): Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J Neurosci* 21:RC159.
Knutson B, Fong GW, Bennett SM, Adams CM, Hommer D (2003): A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. *Neuroimage* 18:263-272.
Knutson B, Westdorp A, Kaiser E, Hommer D (2000): FMRI visualization of brain activity during a monetary incentive delay task. *Neuroimage* 12:20-27.
Kober H, Barrett LF, Joseph J, Bliss-Moreau E, Lindquist K, Wager TD (2008): Functional grouping and cortical-subcortical interactions in emotion: a meta-analysis of neuroimaging studies. *Neuroimage* 42:998-1031.
Koch K, Schachtzabel C, Wagner G, Schikora J, Schultz C, Reichenbach JR, et al (2010): Altered activation in association with reward-related trial-and-error learning in patients with schizophrenia. *Neuroimage* 50:223-232.
Kosaka H, Omori M, Murata T, Iidaka T, Yamada H, Okada T, et al (2002): Differential amygdala response during facial recognition in patients with schizophrenia: an fMRI study. *Schizophr Res* 57:87-95.
Kraepelin E (1971): *Dementia Praecox and Paraphrenia*. Edinburgh: E. & S. Livingstone.
Kring AM, Moran EK (2008): Emotional response deficits in schizophrenia: insights from affective science. *Schizophr Bull* 34:819-834.
Kringelbach ML, Rolls ET (2004): The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Prog Neurobiol* 72:341-372.
Kwapil TR (1998): Social anhedonia as a predictor of the development of schizophrenia-spectrum disorders. *J Abnorm Psychol* 107:558-565.
LaBar KS, Crupain MJ, Voyvodic JT, McCarthy G (2003): Dynamic perception of facial affect and identity in the human brain. *Cereb Cortex* 13:1023-1033.
Lake CR, Hurwitz N (2006): Schizoaffective disorders are psychotic mood disorders; there are no schizoaffective disorders. *Psychiatry Res* 143:255-287.
Lang PJ, Bradley MM, Cuthbert BN (1999): International affective picture system (IAPS): Technical manual and affective ratings. Gainesville: University of Florida.
Lee D, Seo H (2007): Mechanisms of reinforcement learning and decision making in the primate dorsolateral prefrontal cortex. *Ann N Y Acad Sci* 1104:108-122.
Lee J, Park S (2005): Working memory impairments in schizophrenia: a meta-analysis. *J Abnorm Psychol* 114:599-611.
Leifker FR, Bowie CR, Harvey PD (2009): Determinants of everyday outcomes in schizophrenia: the influences of cognitive impairment, functional capacity, and symptoms. *Schizophr Res* 115:82-87.
Locke HS, Braver TS (2008): Motivational influences on cognitive control: behavior, brain activation, and individual differences. *Cogn Affect Behav Neurosci* 8:99-112.
Malhi GS, Green M, Fagiolini A, Peselow ED (2008): Schizoaffective disorder: diagnostic issues and future recommendations. *Bipolar Disord*.
Mamah D, Wang L, Barch D, de Erausquin GA, Gado M, Csernansky JG (2007): Structural analysis of the basal ganglia in schizophrenia. *Schizophr Res* 89:59-71.
Martino DJ, Bucay D, Butman JT, Allegri RF (2007): Neuropsychological frontal impairments and negative symptoms in schizophrenia. *Psychiatry Res* 152:121-128.

Mathews J, Barch DM (2004): Episodic memory for emotional and nonemtional words in schizophrenia. *Cognition and Emotion* 18:721-740.
McAvoy MP, Ollinger JM, Buckner RL (2001): Cluster size thresholds for assessment of significant activation in fMRI. *Neuroimage* 13:S198.
McClure SM, Berns GS, Montague PR (2003): Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38:339-346.
Menon V, Uddin LQ (2010): Saliency, switching, attention and control: a network model of insula function. *Brain Struct Funct* 214:655-667.
Messinger JW, Tremeau F, Antonius D, Mendelsohn E, Prudent V, Stanford AD, et al (2011): Avolition and expressive deficits capture negative symptom phenomenology: implications for DSM-5 and schizophrenia research. *Clin Psychol Rev* 31:161-168.
Miezin FM, Maccotta L, Ollinger JM, Petersen SE, Buckner RL (2000): Characterizing the hemodynamic response: Effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *Neuroimage* 11:735-759.
Milev P, Ho BC, Arndt S, Andreasen NC (2005): Predictive values of neurocognition and negative symptoms on functional outcome in schizophrenia: a longitudinal first-episode study with 7-year follow-up. *Am J Psychiatry* 162:495-506.
Miller EK, Cohen JD (2001): An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 21:167-202.
Minzenberg MJ, Laird AR, Thelen S, Carter CS, Glahn DC (2009): Meta-analysis of 41 functional neuroimaging studies of executive function in schizophrenia. *Arch Gen Psychiatry* 66:811-822.
Montague PR, Dayan P, Sejnowski TJ (1996): A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci* 16:1936-1947.
Moran P, Owen L, Crookes A, Aluzri M, Reveley M (2008): Abnormal prediction error is associated with negative and depressive symptoms in schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry* 32:116-123.
Morris RW, Vercammen A, Lenroot R, Moore L, Langton JM, Short B, et al (2012): Disambiguating ventral striatum fMRI-related bold signal during reward prediction in schizophrenia. *Mol Psychiatry* 17:235, 280-239.
Morris SE, Heerey EA, Gold JM, Holroyd CB (2008): Learning-related changes in brain activity following errors and performance feedback in schizophrenia. *Schizophr Res* 99:274-285.
Mueser KT, McGurk SR (2004): Schizophrenia. *Lancet* 363:2063-2072.
Murray GK, Corlett PR, Clark L, Pessiglione M, Blackwell AD, Honey G, et al (2008): Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Mol Psychiatry* 13:239, 267-276.
Myin-Germeys I, Delespaul PA, deVries MW (2000): Schizophrenia patients are more emotionally active than is assumed based on their behavior. *Schizophr Bull* 26:847-854.
Nielsen MO, Rostrup E, Wulff S, Bak N, Broberg BV, Lublin H, et al (2012): Improvement of Brain Reward Abnormalities by Antipsychotic Monotherapy in Schizophrenia. *Arch Gen Psychiatry*:1-10.
Nitschke JB, Dixon GE, Sarinopoulos I, Short SJ, Cohen JD, Smith EE, et al (2006): Altering expectancy dampens neural response to aversive taste in primary taste cortex. *Nat Neurosci* 9:435-442.
Nuechterlein KH, Subotnik KL, Ventura J, Green MF, Gretchen-Doorly D, Asarnow RF (2012): The puzzle of schizophrenia: tracking the core role of cognitive deficits. *Dev Psychopathol* 24:529-536.
O'Doherty J, Critchley H, Deichmann R, Dolan RJ (2003a): Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. *J Neurosci* 23:7931-7939.
O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ (2004): Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304:452-454.
O'Doherty JP (2004): Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr Opin Neurobiol* 14:769-776.

O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan R (2003b): Temporal difference models and reward-related learning in the human brain. *Neuron* 38:329-337.
O'Doherty JP, Deichmann R, Critchley HD, Dolan RJ (2002): Neural responses during anticipation of a primary taste reward. *Neuron* 33:815-826.
Oades RD (1997): Stimulus dimension shifts in patients with schizophrenia, with and without paranoid hallucinatory symptoms, or obsessive compulsive disorder: strategies, blocking and monoamine status. *Behav Brain Res* 88:115-131.
Oldfield RC (1971): The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9:97-113.
Oorschot M, Lataster T, Thewissen V, Lardinois M, Wichers M, van Os J, et al (2011): Emotional Experience in Negative Symptoms of Schizophrenia--No Evidence for a Generalized Hedonic Deficit. *Schizophr Bull*.
Owen AM, McMillan KM, Laird AR, Bullmore E (2005): N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. *Hum Brain Mapp* 25:46-59.
Pantelis C, Barber FZ, Barnes TR, Nelson HE, Owen AM, Robbins TW (1999): Comparison of set-shifting ability in patients with chronic schizophrenia and frontal lobe damage. *Schizophr Res* 37:251-270.
Paradiso S, Andreasen NC, Crespo-Facorro B, O'Leary DS, Watkins GL, Boles Ponto LL, et al (2003): Emotions in unmedicated patients with schizophrenia during evaluation with positron emission tomography. *Am J Psychiatry* 160:1775-1783.
Peralta V, Cuesta MJ (2008): Exploring the borders of the schizoaffective spectrum: a categorical and dimensional approach. *J Affect Disord* 108:71-86.
Plailly J, d'Amato T, Saoud M, Royet JP (2006): Left temporo-limbic and orbital dysfunction in schizophrenia during odor familiarity and hedonicity judgments. *Neuroimage* 29:302-313.
Poldrack RA, Clark J, Pare-Blagoev EJ, Shohamy D, Creso Moyano J, Myers C, et al (2001): Interactive memory systems in the human brain. *Nature* 414:546-550.
Polli FE, Barton JJ, Thakkar KN, Greve DN, Goff DC, Rauch SL, et al (2008): Reduced error-related activation in two anterior cingulate circuits is related to impaired performance in schizophrenia. *Brain* 131:971-986.
Preacher KJ, Hayes AF (2008): Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav Res Methods* 40:879-891.
Premkumar P, Fannon D, Kuipers E, Simmons A, Frangou S, Kumari V (2008): Emotional decision-making and its dissociable components in schizophrenia and schizoaffective disorder: a behavioural and MRI investigation. *Neuropsychologia* 46:2002-2012.
Rabinowitz J, Levine SZ, Garibaldi G, Bugarski-Kirola D, Berardo CG, Kapur S (2012): Negative symptoms have greater impact on functioning than positive symptoms in schizophrenia: analysis of CATIE data. *Schizophr Res* 137:147-150.
Reske M, Kellermann T, Habel U, Jon Shah N, Backes V, von Wilmsdorff M, et al (2007): Stability of emotional dysfunctions? A long-term fMRI study in first-episode schizophrenia. *J Psychiatr Res* 41:918-927.
Roiser JP, Stephan KE, den Ouden HE, Barnes TR, Friston KJ, Joyce EM (2009): Do patients with schizophrenia exhibit aberrant salience? *Psychol Med* 39:199-209.
Rolls ET, Grabenhorst F, Parris BA (2008a): Warm pleasant feelings in the brain. *Neuroimage* 41:1504-1513.
Rolls ET, Kringelbach ML, de Araujo IE (2003): Different representations of pleasant and unpleasant odours in the human brain. *Eur J Neurosci* 18:695-703.
Rolls ET, McCabe C, Redoute J (2008b): Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. *Cereb Cortex* 18:652-663.

Romaniuk L, Honey GD, King JR, Whalley HC, McIntosh AM, Levita L, et al (2010): Midbrain activation during Pavlovian conditioning and delusional symptoms in schizophrenia. *Arch Gen Psychiatry* 67:1246-1254.

Rossion B, Pourtois G (2004): Revisiting Snodgrass and Vanderwart's object pictorial set: the role of surface detail in basic-level object recognition. *Perception* 33:217-236.

Russell JA (1980): A circumplex model of affect. *J Pers Soc Psychol* 39:1161-1178.

Salamone JD, Correa M, Farrar A, Mingote SM (2007): Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology (Berl)* 191:461-482.

Schlagenhauf F, Sterzer P, Schmack K, Ballmaier M, Rapp M, Wrase J, et al (2009): Reward feedback alterations in unmedicated schizophrenia patients: relevance for delusions. *Biol Psychiatry* 65:1032-1039.

Schlenker R, Cohen R, Hopmann G (1995): Affective modulation of the startle reflex in schizophrenic patients. *Eur Arch Psychiatry Clin Neurosci* 245:309-318.

Schneider F, Habel U, Reske M, Toni I, Falkai P, Shah NJ (2007): Neural substrates of olfactory processing in schizophrenia patients and their healthy relatives. *Psychiatry Res* 155:103-112.

Schneider F, Weiss U, Kessler C, Salloum JB, Posse S, Grodd W, et al (1998): Differential amygdala activation in schizophrenia during sadness. *Schizophr Res* 34:133-142.

Schultz W (1997): Dopamine neurons and their role in reward mechanisms. *Curr Opin Neurobiol* 7:191-197.

Schultz W (2001): Reward signaling by dopamine neurons. *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry* 7:293-302.

Schultz W (2007): Multiple dopamine functions at different time courses. *Annu Rev Neurosci* 30:259-288.

Schultz W, Apicella P, Ljungberg T (1993): Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *The Journal of Neuroscience* 13:900-913.

Schultz W, Dayan P, Montague PR (1997): A neural substrate of prediction and reward. *Science* 275:1593-1599.

Schwarz G (1978): Estimating the dimension of a model. *Annals of Statistics* 6:461-464.

Seiferth NY, Pauly K, Habel U, Kellermann T, Jon Shah N, Ruhrmann S, et al (2008): Increased neural response related to neutral faces in individuals at risk for psychosis. *Neuroimage* 40:289-297.

Sevy S, Burdick KE, Visweswaraiah H, Abdelmessih S, Lukin M, Yechiam E, et al (2007): Iowa gambling task in schizophrenia: a review and new data in patients with schizophrenia and co-occurring cannabis use disorders. *Schizophr Res* 92:74-84.

Shoichet RP, Oakley A (1978): Notes on the treatment of anhedonia. *Can Psychiatr Assoc J* 23:487-492.

Simon JJ, Biller A, Walther S, Roesch-Ely D, Stippich C, Weisbrod M, et al (2010): Neural correlates of reward processing in schizophrenia--relationship to apathy and depression. *Schizophr Res* 118:154-161.

Smith KS, Berridge KC (2007): Opioid limbic circuit for reward: interaction between hedonic hotspots of nucleus accumbens and ventral pallidum. *J Neurosci* 27:1594-1605.

Snaith RP, Hamilton M, Morley S, Humayan A, Hargreaves D, Trigwell P (1995): A scale for the assessment of hedonic tone the Snaith-Hamilton Pleasure Scale. *Br J Psychiatry* 167:99-103.

Snyder AZ (1996): Difference image versus ratio image error function forms in PET-PET realignment. In: Bailer D, Jones T editors. *Quantification of Brain Function Using PET*. San Diego: Academic Press.

Somlai Z, Moustafa AA, Keri S, Myers CE, Gluck MA (2011): General functioning predicts reward and punishment learning in schizophrenia. *Schizophr Res* 127:131-136.

Starkstein SE, Mayberg HS, Preziosi TJ, Andrezejewski P, Leiguarda R, Robinson RG (1992): Reliability, validity, and clinical correlates of apathy in Parkinson's disease. *J Neuropsychiatry Clin Neurosci* 4:134-139.

Strauss GP, Gold JM (2012): A new perspective on anhedonia in schizophrenia. *Am J Psychiatry* 169:364-373.

Surguladze S, Russell T, Kucharska-Pietura K, Travis MJ, Giampietro VP, David AS, et al (2006): A reversal of the normal pattern of parahippocampal response to neutral and fearful faces is associated with reality distortion in schizophrenia. *Biol Psychiatry* 60:423-431.

Sutton R (1988): Learning to predict by the method of temporal difference. *Machine Learning* 3:9-44.

Takahashi H, Koeda M, Oda K, Matsuda T, Matsushima E, Matsuura M, et al (2004): An fMRI study of differential neural response to affective pictures in schizophrenia. *Neuroimage* 22:1247-1254.

Talairach J, Tournoux P (1988): *Co-planar stereotaxic atlas of the human brain*. New York: Thieme.

Taylor SF, Kang J, Brege IS, Tso IF, Hosanagar A, Johnson TD (2012): Meta-analysis of functional neuroimaging studies of emotion perception and experience in schizophrenia. *Biol Psychiatry* 71:136-145.

Taylor SF, Liberzon I, Decker LR, Koeppe RA (2002): A functional anatomic study of emotion in schizophrenia. *Schizophr Res* 58:159-172.

Taylor SF, Phan KL, Britton JC, Liberzon I (2005): Neural response to emotional salience in schizophrenia. *Neuropsychopharmacology* 30:984-995.

Treadway MT, Zald DH (2011): Reconsidering anhedonia in depression: lessons from translational neuroscience. *Neurosci Biobehav Rev* 35:537-555.

Trémeau F (2006): A review of emotion deficits in schizophrenia. *Dialogues in clinical neuroscience* 8:59-70.

Turnbull OH, Evans CE, Kemish K, Park S, Bowman CH (2006): A novel set-shifting modification of the iowa gambling task: flexible emotion-based learning in schizophrenia. *Neuropsychology* 20:290-298.

Turner DC, Clark L, Pomarol-Clotet E, McKenna P, Robbins TW, Sahakian BJ (2004): Modafinil improves cognition and attentional set shifting in patients with chronic schizophrenia. *Neuropsychopharmacology* 29:1363-1373.

Tyson PJ, Laws KR, Roberts KH, Mortimer AM (2004): Stability of set-shifting and planning abilities in patients with schizophrenia. *Psychiatry Res* 129:229-239.

Ursu S, Kring AM, Gard MG, Minzenberg MJ, Yoon JH, Ragland JD, et al (2011): Prefrontal cortical deficits and impaired cognition-emotion interactions in schizophrenia. *Am J Psychiatry* 168:276-285.

Vollema MG, van den Bosch RJ (1995): The multidimensionality of schizotypy. *Schizophr Bull* 21:1.

Wacker J, Dillon DG, Pizzagalli DA (2009): The role of the nucleus accumbens and rostral anterior cingulate cortex in anhedonia: integration of resting EEG, fMRI, and volumetric techniques. *Neuroimage* 46:327-337.

Wager TD, Barrett, L.F., Bliss-Moreau, E., Lindquist, K., Duncan, S., Kober, H., Joseph, J., Davidson, M., Mize, J. (2008): The Neuroimaging of Emotion. In: Lewis M, Haviland-Jones, J.M., Barrett, L.F. editor. *Handbook of Emotion, 3rd Edition*. New York: Guilford, pp 249-271.

Wager TD, Smith EE (2003): Neuroimaging studies of working memory: a meta-analysis. *Cogn Affect Behav Neurosci* 3:255-274.

Wallis JD (2007): Orbitofrontal cortex and its contribution to decision-making. *Annu Rev Neurosci* 30:31-56.

Walter H, Heckers S, Kassubek J, Erk S, Frasch K, Abler B (2010): Further evidence for aberrant prefrontal salience coding in schizophrenia. *Front Behav Neurosci* 3:62.

Walter H, Kammerer H, Frasch K, Spitzer M, Abler B (2009): Altered reward functions in patients on atypical antipsychotic medication in line with the revised dopamine hypothesis of schizophrenia. *Psychopharmacology (Berl)* 206:121-132.

Waltz JA, Frank MJ, Robinson BM, Gold JM (2007): Selective reinforcement learning deficits in schizophrenia support predictions from computational models of striatal-cortical dysfunction. *Biol Psychiatry* 62:756-764.

Waltz JA, Frank MJ, Wiecki TV, Gold JM (2011): Altered probabilistic learning and response biases in schizophrenia: behavioral evidence and neurocomputational modeling. *Neuropsychology* 25:86-97.

Waltz JA, Gold JM (2007): Probabilistic reversal learning impairments in schizophrenia: further evidence of orbitofrontal dysfunction. *Schizophr Res* 93:296-303.

Waltz JA, Schweitzer JB, Gold JM, Kurup PK, Ross TJ, Salmeron BJ, et al (2009): Patients with schizophrenia have a reduced neural response to both unpredictable and predictable primary reinforcers. *Neuropsychopharmacology* 34:1567-1577.

Waltz JA, Schweitzer JB, Ross TJ, Kurup PK, Salmeron BJ, Rose EJ, et al (2010): Abnormal responses to monetary outcomes in cortex, but not in the basal ganglia, in schizophrenia. *Neuropsychopharmacology* 35:2427-2439.

Wang L, Mamah D, Harms MP, Karnik M, Price JL, Gado MH, et al (2008): Progressive deformation of deep brain nuclei and hippocampal-amygdala formation in schizophrenia. *Biol Psychiatry* 64:1060-1068.

Weickert TW, Goldberg TE, Callicott JH, Chen Q, Apud JA, Das S, et al (2009): Neural correlates of probabilistic category learning in patients with schizophrenia. *J Neurosci* 29:1244-1254.

Weickert TW, Goldberg TE, Egan MF, Apud JA, Meeter M, Myers CE, et al (2010): Relative risk of probabilistic category learning deficits in patients with schizophrenia and their siblings. *Biol Psychiatry* 67:948-955.

Weickert TW, Terrazas A, Bigelow LB, Malley JD, Hyde T, Egan MF, et al (2002): Habit and skill learning in schizophrenia: evidence of normal striatal processing with abnormal cortical input. *Learn Mem* 9:430-442.

Weiler JA, Bellebaum C, Brune M, Juckel G, Daum I (2009): Impairment of probabilistic reward-based learning in schizophrenia. *Neuropsychology* 23:571-580.

Williams LM, Das P, Harris AW, Liddell BB, Brammer MJ, Olivieri G, et al (2004): Dysregulation of arousal and amygdala-prefrontal systems in paranoid schizophrenia. *Am J Psychiatry* 161:480-489.

Wise RA (1982): Neuroleptics and operant behavior: The anhedonia hypothesis. *Behav Brain Sci* 5:39-87.

Woods RP, Cherry SR, Mazziotta JC (1992): Rapid automated algorithm for aligning and reslicing PET images. *J Comput Assist Tomogr* 16:620-633.

Worsley KJ (1994): Local Maxima and the expected Euler Characteristic of Excursion Sets of chi-squared F and t Fields. *Adv Appl Prob* 26:13-42.

Yilmaz A, Simsek F, Gonul AS (2012): Reduced reward-related probability learning in schizophrenia patients. *Neuropsychiatr Dis Treat* 8:27-34.

Yip SW, Sacco KA, George TP, Potenza MN (2009): Risk/reward decision-making in schizophrenia: a preliminary examination of the influence of tobacco smoking and relationship to Wisconsin Card Sorting Task performance. *Schizophr Res* 110:156-164.

Zald DH (2003): The human amygdala and the emotional evaluation of sensory stimuli. *Brain Research Reviews* 41:88-123.

Ziauddeen H, Murray GK (2010): The relevance of reward pathways for schizophrenia. *Curr Opin Psychiatry* 23:91-96.

Zink CF, Pagnoni G, Martin-Skurski ME, Chappelow JC, Berns GS (2004): Human striatal responses to monetary reward depend on saliency. *Neuron* 42:509-517.